

PCT

ORGANISATION MONDIALE DE LA PROPRIÉTÉ INTELLECTUELLE
Bureau international

DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPERATION EN MATIÈRE DE BREVETS (PCT)

(51) Classification internationale des brevets ⁶ : C07K 14/00	A2	(11) Numéro de publication internationale: WO 99/67282 (43) Date de publication internationale: 29 décembre 1999 (29.12.99)
(21) Numéro de la demande internationale: PCT/FR99/01495 (22) Date de dépôt international: 22 juin 1999 (22.06.99) (30) Données relatives à la priorité: 98/08094 25 juin 1998 (25.06.98) FR (71) Déposant (pour tous les Etats désignés sauf US): CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) [FR/FR]; 3, rue Michel Ange, F-75794 Paris Cedex 16 (FR). (72) Inventeurs; et (75) Inventeurs/Déposants (US seulement): CLEMENT, Jean-Luc [FR/FR]; 550, rue Paradis, F-13008 Marseille (FR). TIRARD, Alain [FR/FR]; Chemin de la Floride, F-83640 Saint-Zacharie (FR). RENUCCI, Marielle [FR/FR]; 11, rue Dumont d'Urville, F-13008 Marseille (FR). BELAICH, Anne [FR/FR]; Villa 8, 33, rue Floralia, F-13009 Marseille (FR). MATARAZZO, Valéry [FR/FR]; 4, rue du Calcaire, F-13360 Roquevaire (FR). (74) Mandataires: BREESE, Pierre etc.; Breese-Majerowicz, 3, avenue de l'Opéra, F-75001 Paris (FR).		(81) Etats désignés: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, brevet ARIPO (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), brevet eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Publiée <i>Sans rapport de recherche internationale, sera republiée dès réception de ce rapport.</i>
(54) Title: NOVEL OLFACTORY RECEPTORS AND THEIR USES (54) Titre: NOUVEAUX RECEPTEURS OLFACTIFS ET LEURS UTILISATIONS (57) Abstract <p>The invention concerns novel olfactory receptors, and the genes coding for said receptors. The invention also concerns the use of said receptors for detecting aromas, quality control, sample analysis or comparing perfumes, detecting toxic substances or trapping smells.</p> (57) Abrégé <p>La présente invention concerne de nouveaux récepteurs olfactifs, ainsi que les gènes codant ces récepteurs. L'invention se rapporte aussi à l'utilisation de ces récepteurs pour la détection d'arômes, le contrôle de qualité, l'analyse d'échantillons, l'analyse ou la comparaison de parfums, la détection de substances toxiques, ou le piégeage d'odeurs.</p>		

UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AL	Albanie	ES	Espagne	LS	Lesotho	SI	Slovénie
AM	Arménie	FI	Finlande	LT	Lituanie	SK	Slovaquie
AT	Autriche	FR	France	LU	Luxembourg	SN	Sénégal
AU	Australie	GA	Gabon	LV	Lettonie	SZ	Swaziland
AZ	Azerbaïdjan	GB	Royaume-Uni	MC	Monaco	TD	Tchad
BA	Bosnie-Herzégovine	GE	Géorgie	MD	République de Moldova	TG	Togo
BB	Barbade	GH	Ghana	MG	Madagascar	TJ	Tadjikistan
BE	Belgique	GN	Guinée	MK	Ex-République yougoslave de Macédoine	TM	Turkménistan
BF	Burkina Faso	GR	Grèce	ML	Mali	TR	Turquie
BG	Bulgarie	HU	Hongrie	MN	Mongolie	TT	Trinité-et-Tobago
BJ	Bénin	IE	Irlande	MR	Mauritanie	UA	Ukraine
BR	Brésil	IL	Israël	MW	Malawi	UG	Ouganda
BY	Bélarus	IS	Islande	MX	Mexique	US	Etats-Unis d'Amérique
CA	Canada	IT	Italie	NE	Niger	UZ	Ouzbékistan
CF	République centrafricaine	JP	Japon	NL	Pays-Bas	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norvège	YU	Yougoslavie
CH	Suisse	KG	Kirghizistan	NZ	Nouvelle-Zélande	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	République populaire démocratique de Corée	PL	Pologne		
CM	Cameroun	KR	République de Corée	PT	Portugal		
CN	Chine	KZ	Kazakstan	RO	Roumanie		
CU	Cuba	LC	Sainte-Lucie	RU	Fédération de Russie		
CZ	République tchèque	LI	Liechtenstein	SD	Soudan		
DE	Allemagne	LK	Sri Lanka	SE	Suède		
DK	Danemark	LR	Libéria	SG	Singapour		
EE	Estonie						

NOUVEAUX RÉCEPTEURS OLFACTIFS ET LEURS
UTILISATIONS.

5 La présente invention se rapporte à la mise
en évidence de nouveaux récepteurs olfactifs chez la
marmotte, au clonage et au séquençage des gènes codant
pour ces récepteurs ainsi qu'à l'utilisation de ceux-ci
pour le criblage de ligands et la préparation de
10 biocapteurs.

 La découverte récente de récepteurs
olfactifs de vertébrés bouleverse les stratégies
initialement envisagées pour la conception et la
réalisation de nez artificiel avec des capteurs physico-
15 chimiques. En effet, au début des années 1990, les
biologistes sont parvenus à isoler à partir d'épithélium
olfactif de mammifères et à séquencer les premières
protéines constituant des récepteurs olfactifs (3), et
c'est en 1993 que le premier récepteur olfactif a été
20 exprimé (7). Or, il est admis que l'homme, qui est
réputé avoir un odorat limité, est capable de
différencier plus de 10 000 molécules odorantes et que
1% de son génome est composé de gènes codant pour les
récepteurs olfactifs (1). On conçoit donc le champ
25 d'investigation formidable qui s'ouvre aux chercheurs en
matière de capteurs biologiques potentiels. En outre, il
apparaît déjà que ces capteurs biologiques ont une
sensibilité supérieure d'environ 100 000 fois aux
meilleurs capteurs physico-chimiques existants (4, 6).
30 Des travaux plus récents ont montré que ces détecteurs
sont également sensibles à des molécules non biologiques
(5).

 Tous les organismes vivants dépendent
35 d'informations sensorielles pour leur survie. Les
perceptions sensorielles sont transmises par les organes

des sens qui reçoivent des stimuli physiques (vue, ouïe, toucher) et chimiques (gout, odorat). Dans la plupart des espèces, la perception des stimuli chimiques est essentielle pour l'accomplissement de plusieurs tâches vitales comme localiser la nourriture, identifier des partenaires, identifier la progéniture et détecter les prédateurs ou d'autres dangers. Dans certaines espèces, l'odorat permet aussi la communication, à des distances pouvant atteindre plusieurs kilomètres, entre individus et permettre de rassembler le groupe, les réactions d'attaque et de défense, l'activité de reproduction et l'allaitement. Les molécules odorantes peuvent aussi induire des changements physiologiques.

Dans la plupart des cas les odeurs résultent d'une combinaison complexe de plusieurs molécules. Cette complexité soulève d'intéressantes questions sur les caractéristiques des récepteurs pour permettre aux animaux de reconnaître une myriade de molécules odorantes (estimées à plus de 10 000) à des concentrations aussi basses que 10^{-12} M. Il semble que la reconnaissance est basée sur une large famille multigénique de récepteurs d'odeurs comprenant plusieurs centaines ou milliers de sous types. Ces récepteurs sont supposés contenir 7 domaines trans-membranaires, à partir de l'hypothèse selon laquelle les signaux odorants sont transduits par des cascades de réactions couplées aux protéines G dans les neurones olfactifs sensitifs. La transduction résulte en une augmentation de seconds messagers comme les nucléotides cycliques ou l'inositol triphosphate et à leur tour ces messagers activent les canaux iono-dépendants et la phosphorylation de plusieurs protéines parmi lesquelles les récepteurs d'odeurs eux-mêmes.

Buck et Axel (3) ont d'abord caractérisé des récepteurs d'odeurs de rat à l'aide de techniques d'amplification (PCR) et d'amorces dégénérées

correspondant aux domaines les plus conservés des récepteurs couplés aux protéines G. Depuis ces premiers travaux, plus de 339 récepteurs ont été séquencés, le plus souvent partiellement, parmi une grande variété d'espèces dont l'homme, le chien, la souris, le poulet, deux espèces de poisson, deux espèces d'amphibiens et un nématode. Beaucoup d'espèces restent cependant encore à étudier et on estime que plus de 1000 gènes (soit 1% du génome) codent pour la superfamille des récepteurs olfactifs. Les mécanismes sous-jacents à la perception olfactive sont singuliers et uniques en comparaison avec d'autres systèmes sensoriels et une étude plus poussée dans ce domaine qui a d'importantes implications pour l'identification de ces protéines est nécessaire.

Plusieurs travaux ont souligné l'importance de l'olfaction pour la marmotte des Alpes (2). Des études éthologiques et analytiques ont montré qu'un groupe de 40 composés, produits par les glandes jugales, était utilisés pour marquer le territoire et identifier le groupe social. Les travaux réalisés dans le cadre de la présente invention sur l'épithélium olfactif de la marmotte des Alpes visaient notamment à obtenir un nombre suffisant de séquences de récepteurs olfactifs pour permettre une comparaison significative avec les séquences de vertébrés précédemment déterminées. Une stratégie basée sur la RT-PCR a été utilisée pour identifier des séquences putatives de récepteurs olfactifs de marmotte. Des oligonucléotides dégénérés correspondant à la séquence des domaines conservés dans le second domaine transmembranaire, la seconde boucle intracellulaire et le 7ème domaine transmembranaire de récepteurs olfactifs ont été utilisés par paires comme amorces pour des PCR à partir de l'ADN complémentaire obtenu en utilisant l'ARNm de l'épithélium nasal de marmotte.

Les travaux de recherche réalisés dans le cadre de la présente invention ont ainsi permis pour la

première fois d'identifier, de cloner et de séquencer de nouveaux récepteurs olfactifs chez la marmotte. Ces récepteurs sont utiles pour la conception et la mise au point de biocapteurs ou la préparation de cellules transfectées. Ainsi, ces récepteurs peuvent être associés à des membranes artificielles qui seront utilisées dans différents biocapteurs disposés en parallèle, chacun possédant un type particulier de récepteur, l'ensemble étant géré par un logiciel de réseaux de neurones formels pour constituer un système de détection de type nez électronique dont les capteurs sont des capteurs bioélectroniques.

L'invention se rapporte donc à un récepteur olfactif purifié de marmotte.

La distinction entre des dizaines de milliers d'odeurs dépend d'une myriade de récepteurs situés à la surface des dendrites neuronales de l'épithélium nasal. En utilisant l'épithélium nasal de la marmottes des Alpes et différents jeux d'amorces dégénérées correspondant à des séquences consensus de récepteurs d'odeurs, les inventeurs sont parvenus à amplifier par PCR-inverse (RT-PCR), cloner et obtenir la séquence partielle de 23 nouveaux produits de gènes codant pour des récepteurs d'odeurs. Après consultation par le programme blast des banques de données du NCBI, leur traduction en séquences d'acides aminés montre une forte similarité avec des séquences protéiques de récepteurs d'odeurs précédemment rapportées, et les classe sans ambiguïté dans la même superfamille de récepteurs à 7 domaines transmembranaires. Les régions hélicoïdales transmembranaires III, IV et V, ainsi que les boucles intra et extracellulaires ont été définies par l'établissement d'un profil d'hydropathie et la prédiction par ordinateur de la structure secondaire. Dans une première tentative de cartographie des sites de fixation d'odeur, les inventeurs ont réalisé une analyse

de variabilité du type de celle décrite par Wu et Kabat (8) sur les régions déterminant la complémentarité (CDR) des immunoglobulines. Quatre pics principaux de variabilité sont localisés à l'intérieur des 1ère et 3ème boucles extracellulaires prédites, et à l'intérieur des 4ème et 5ème domaines transmembranaires prédits. Ces positions feraient donc partie du site spécifique de liaison des molécules odorantes. Les comparaisons avec la séquence de récepteurs olfactifs d'autres espèces suggèrent que les séquences de la marmotte déterminées dans cette étude appartiennent à trois familles différentes.

L'invention concerne donc plus particulièrement un récepteur olfactif purifié constitué par ou comprenant la séquence en acides aminés choisie parmi celles représentées dans la liste de séquences en annexe sous les numéros SEQ ID No:1 à SEQ ID No:23, ou un dérivé fonctionnellement équivalent de celles-ci. On entend par dérivé équivalent de ces séquences, les séquences comprenant une modification et/ou une suppression et/ou une addition d'un ou plusieurs résidus d'acides aminés mais conservant environ 75% et de préférence au moins 95% d'homologie avec la séquence dont elle est dérivée. Les récepteurs de l'invention présentent des régions très conservées et des régions très hétérogènes. On considère que les régions très conservées sont celles conférant à la protéine son caractère de récepteur, alors que les régions très hétérogènes sont celles qui confèrent à chaque récepteur sa spécificité. Ainsi, selon l'application envisagée, il est possible de préparer des dérivés des récepteurs de l'invention dont la spécificité est modifiée mais qui restent dans le cadre de la présente invention.

L'invention a également pour objet les anticorps poly ou monoclonaux dirigés contre au moins un

récepteur de l'invention, un dérivé ou un fragment de ceux-ci. Ces anticorps peuvent être préparés par les méthodes décrites dans la littérature. Les anticorps polyclonaux sont formés selon les techniques classiques par injection des protéines, extraites à partir d'épithélium ou produite par transformation génétique d'un hôte, à des animaux, puis récupération des antisérums et des anticorps à partir des antisérums par exemple par chromatographie d'affinité. Les anticorps monoclonaux peuvent être produits en fusionnant des cellules de myélomes avec des cellules de rates d'animaux préalablement immunisés à l'aide des récepteurs de l'invention. Ces anticorps sont utiles pour rechercher de nouveaux récepteurs olfactifs ou les homologues de ces récepteurs chez d'autres mammifères ou encore pour étudier la parenté entre des récepteurs de différents individus ou espèces.

L'invention se rapporte également à une molécule d'acide nucléique comprenant ou constituée par une séquence nucléique codant pour un récepteur tel que défini précédemment. Plus particulièrement, l'invention se rapporte à une molécule d'acide nucléique comprenant ou constituée par une séquence choisie parmi celles représentées dans la liste de séquences en annexe sous les numéro SEQ ID No:24 à SEQ ID No:47, lesquelles codent respectivement pour les récepteurs dont les séquences en acides aminés sont représentées dans la liste de séquences en annexe sous les numéro SEQ ID No:1 à SEQ ID No:23.

L'invention concerne bien entendu aussi les séquences nucléotidiques dérivées des séquences ci-dessus, par exemple du fait de la dégénérescence du code génétique, et qui code pour des protéines présentant des caractéristiques et propriétés de récepteurs olfactifs.

L'invention concerne également un vecteur comprenant au moins une molécule d'acide nucléique précédente, avantageusement associée à des séquences de contrôle adaptés, ainsi qu'un procédé de production ou d'expression dans un hôte cellulaire d'un récepteur de l'invention ou d'un fragment de celui-ci. La préparation de ces vecteurs ainsi que la production ou l'expression dans un hôte des protéines de l'invention peuvent être réalisées par les techniques de biologie moléculaire et de génie génétique bien connues de l'homme du métier.

A titre d'exemple, un procédé de production d'un récepteur selon l'invention consiste :

- à transférer une molécule d'acide nucléique de l'invention ou un vecteur contenant ladite molécule dans un hôte cellulaire,
- à cultiver ledit hôte cellulaire dans des conditions permettant la production de la protéine constituant le récepteur,
- à isoler, par tous moyens appropriés les dites protéines.

A titre d'exemple, un procédé d'expression d'un récepteur selon l'invention consiste :

- à transférer une molécule d'acide nucléique de l'invention ou un vecteur contenant ladite molécule dans un hôte cellulaire,
- à cultiver ledit hôte cellulaire dans des conditions permettant l'expression desdites récepteurs à la surface de l'hôte.

L'hôte cellulaire mis en oeuvre dans les procédés précédents peut être choisi parmi les procaryotes ou les eucaryotes et notamment parmi les bactéries, les levures, les cellules de mammifères, de plantes ou d'insectes.

L'expression dans des cellules eucaryotes est préférable pour que les récepteurs puissent subir

les modifications post-traductionnelles nécessaires à leur fonction.

Une molécule d'acide nucléique codant pour un récepteur olfactif ou un vecteur selon l'invention
5 peuvent aussi être utilisés pour transformer des animaux et établir une lignée d'animaux transgéniques.

Le vecteur utilisé est choisi en fonction de l'hôte dans lequel il sera transféré; il peut s'agir de
10 tout vecteur comme un plasmide.

L'invention concerne donc aussi les hôtes cellulaires exprimant des récepteurs olfactifs obtenus conformément aux procédés précédents.

L'invention concernent également les sondes nucléiques et oligonucléotides préparés à partir des
15 molécules d'acide nucléique de l'invention.

Ces sondes, avantageusement marquées, sont utiles pour la détection par hybridation de séquences
similaires de récepteurs chez d'autres individus ou
20 espèces. Selon les techniques classiques, ces sondes sont mises en contact avec un échantillon biologique. Différentes techniques d'hybridation peuvent être mises en oeuvre telles que l'hybridation sur taches (Dot-blot) ou l'hybridation sur répliques (technique de Southern)
25 ou autres techniques (DNA chips). De telles sondes constituent des outils permettant de détecter rapidement des séquences similaires dans les gènes codant pour des récepteurs olfactifs ce qui permet d'étudier la présence, l'origine et la conservation de ces protéines.

Les oligonucléotides sont utiles pour des
30 expériences de PCR par exemple pour rechercher des gènes dans d'autres espèces ou dans un but de diagnostic.

Comme indiqué précédemment, les récepteurs
35 olfactifs sont des protéines à 7 domaines transmembranaires couplées aux protéines G. La fixation

d'un ligand sur le récepteur entraîne un changement de conformation du récepteur, et à l'intérieur de la cellule, ce signal est transduit par l'intermédiaire de seconds messagers. En conséquence, l'invention a pour
5 objet un procédé de criblage de composés susceptibles de constituer des ligands des récepteurs décrits précédemment consistant à mettre en contact un composé et un ou plusieurs desdits récepteurs et à mesurer par tout moyen approprié l'affinité entre ledit composé et
10 ledit récepteur.

La mise en contact entre le composé à tester et le ou les récepteurs olfactifs de l'invention peut être réalisée en utilisant des hôtes décrits précédemment et exprimant à leur surface au moins
15 desdits récepteurs. Il peut s'agir d'une lignée de cellules immortalisées olfactives ou non, transfectée par un vecteur portant l'ADNc permettant d'exprimer à sa surface et à un niveau élevé un récepteur olfactif recombinant fonctionnel. Si le composé testé constitue
20 un ligand, sa mise en contact avec des cellules transformées, induit des signaux intracellulaires qui découlent de la fixation dudit composé sur le récepteur.

La mise en contact des composés à tester avec les récepteurs de l'invention peut aussi être
25 réalisée en fixant un ou plusieurs récepteurs sur une ou plusieurs membranes. Les récepteurs olfactifs de l'invention peuvent donc aussi être intégrés à un biocapteur. Dans un tel système, il est possible de visualiser en temps réel des interactions entre le
30 composé testé et le récepteur. L'un des partenaires du couple récepteur/ligand est fixé sur une interface qui peut contenir une matrice recouverte de chaînes aliphatiques. Cette matrice hydrophobe peut être facilement recouverte d'une couche lipidique par fusion
35 spontanée de liposomes injectés à son contact. Des récepteurs olfactifs insérés dans des liposomes ou des

vésicules peuvent ainsi être intégrés à des biocapteurs. Les ligands sont ainsi analysés vis-à-vis d'un ou plusieurs récepteurs olfactifs différents.

5 Les méthodes ci-dessus permettent de déterminer si un composé active ou inhibe les récepteurs. Dans ce mode de réalisation, il est avantageux de disposer d'un ligand connu qui permet des mesures par compétition.

10 L'invention se rapporte donc aussi à un composé non encore connu constituant un ligand d'un récepteur olfactif, identifié et sélectionné par le procédé ci-dessus.

15 Les récepteurs de l'invention trouvent des applications dans des domaines très variés comme :

- L'industrie agroalimentaire, pour la détection d'arômes, le contrôle de qualité, l'analyse d'échantillons.

20 - La parfumerie, pour l'analyse ou la comparaison de parfums.

- L'environnement, pour la détection de substances toxiques, comme des gaz ou le piégeage d'odeurs.

25 D'autres avantages et caractéristiques de l'invention apparaîtront à la lecture des exemples qui suivent concernant l'identification et le clonage des récepteurs olfactifs de marmotte, et qui se réfèrent aux dessins en annexe dans lesquels :

30 - La figure 1 représente l'analyse des produits de PCR réalisée à partir de deux types d'ADNc (R et T) et 3 jeux d'amorces (c-t, 4-1 et 3-2). Les produits de la réaction ont été analysés par électrophorèse sur un gel d'agarose 2 %, comme décrit
35 ci-après dans matériel et méthode. La taille des fragments a été estimée par comparaison avec un standard

de taille connue (côté droit). Les dépôts dans les pistes marquées d'une astérisque contiennent les fragments de la taille attendue.

5 - La figure 2 représente l'alignement de 14 des 23 séquences de récepteurs olfactifs putatifs de marmotte. 14 séquences différentes (AMOR 1 à AMOR 14) ont été analysées à l'aide du logiciel Clustalw. Les régions ombrées indiquent les domaines consensus contenant les acides aminés presque (.) ou totalement (*) conservés. Les domaines transmembranaires (DII à DVII), les boucles extracellulaires (E1 à E3) et les boucles intracellulaires (i2 à i3) ont été délimitées après la détermination des domaines hydrophobes.

15 - La figure 3 représente les profils d'hydropathie des séquences longues obtenues avec le jeu d'amorces c-t (AMOR 1 à AMOR 7) et les séquences courtes obtenues avec le jeu d'amorces 3-2 (AMOR 8 à AMOR 14) ont été obtenues comme décrit dans Matériel et Méthodes. Les séquences longues contiennent 6 régions de forte hydrophobicité (pics) séparées par 5 vallées plus hydrophiles. Les séquences courtes présentent seulement 4 régions de forte hydrophobicité et 3 régions hydrophiles. Ces graphes sont compatibles avec la présence de 6 ou 4 domaines transmembranaires, pour les séquences longues et courtes respectivement. Cette architecture est confirmée par les prédictions d'hélices transmembranaires du programme PHD.

25 La figure 4 représente l'analyse de la variabilité des 14 nouvelles séquences ininterrompues de récepteur olfactif de marmotte. Graphe du haut : variabilité des résidus calculée pour l'alignement de la figure 2. La localisation des pics (positions les plus variables) et la forme globale de la courbe sont indépendantes de la formule utilisée (Wu & Kabat, 30 complexité ou nombre de résidus pris en compte). Graphe du bas : index moyen d'hydropathie des séquences

alignées. Les pics correspondent aux régions hydrophiles (boucles) et les vallées aux régions hydrophobes (domaines transmembranaires). Le graphe minimise l'hydrophobicité du fragment 1 à 59 car la moitié des séquences manque à ces positions. Alors que la position 210 illustre la variabilité habituelle des boucles hydrophiles exposées, la position 148 présente la forte variabilité la plus surprenante dans une région fortement hydrophobe (hélicoïdale) de la molécule.

La figure 5 représente un dendrogramme montrant les similarités entre des récepteurs olfactifs de différentes espèces. Les séquences de récepteurs olfactifs d'autres espèces proviennent de la banque de données NCBI. Il y a cinq familles (notées à gauche). Les astérisques indiquent les séquences pour lesquelles le pourcentage de similarité entre espèce excède 70 %. Abréviations : H : homme ; F : poisson ; C : poulet ; N : nématode ; B : abeille ; A : amphibiens ; D : chien ; M : souris et MM : marmotte.

I - MATÉRIEL ET MÉTHODES.

1) Préparation des tissus.

L'épithélium olfactif a été prélevé sur une marmotte sauvage morte. Pendant la dissection, la tête a été maintenue congelée dans de la carboglace. Les tissus ont été gardés à -80°C jusqu'à leur utilisation.

2) Isolement de l'ARNm.

Les tissus congelés ont été réduits en poussière en les concassant avec un pilon dans un mortier. Le pilon et le mortier étaient refroidis dans de la carboglace et tout l'équipement était stérile. L'ARNm poly(A)+ a été isolé en utilisant le Micro-Fast Track Kit (Invitrogen) puis testé avec le DNA DipStick Kit (Invitrogen).

3) Transcription de l'ADNc.

L'ARNm poly(A)+ a été transcrit en ADNc à l'aide d'une transcriptase inverse puis amplifié par PCR. Pour augmenter la production du premier brin d'ADNc complet le kit cDNA Cycle Kit a été utilisé. La transcription inverse a été faite à partir de 150 ng d'ARNm poly(A)+ en utilisant des amorces oligo dT ou des random primers. Après extraction avec phénol / H2O / EDTA (v/v/v : 1/20/80), l'ADNc de la phase aqueuse a été précipité en présence d'acétate d'ammonium et de glycogène entraîneur dans de l'éthanol glacial à -80°C.

4) PCR.

Trois jeux d'oligonucléotides dégénérés spécifiques de récepteurs olfactif sont été synthétisés pour amplifier ces récepteurs de marmotte.

A partir de résultats antérieurs obtenus chez le rat (3), deux jeux d'amorces ont été synthétisées contre des régions conservées du second et du septième domaine transmembranaire des récepteurs olfactifs.

Amorce 4 : 5'-CC(CT) ATG TA(TC) TTI TT(TC) CT (CT) I(GC)(CT) AA(TC) (TC)TI TC.

Amorce C : 5'-CC(CT) ATG TA(TC) TTG TT(TC) CT(CT) G(GC)(CT) AA(TC) (TC)TG TC-.

Amorce 1 : 5'-(AG)TT (TC)C(TG) IA(AG) (AG)(CG)(AT) (AG)TA IAT (GA)A(AT) IGG (AG)TT.

Amorce T : 5'-GCA CTG CAG AT(AG) AAI GG(AG) TTI A(AG) ATI GG.

Ces combinaisons d'amorces ont été conçues pour permettre d'amplifier des produits de l'ordre de 720 pb.

A partir de résultats antérieurs obtenus chez le rat (3) et le poisson-chat, le 3ème jeu d'oligonucléotides dégénérés a été synthétisé à partir

des régions conservées de la 2ème boucle intracellulaire et du 7ème domaine transmembranaire.

Amorce 3 : 5'-CAC AAG CTT TIG CIT A(TC)G A(CT)A G(AG)T (TA)(TC)(TCG) TIG C.

5 Amorce 2 : 5'-GCA CTG CAG AT(AG) AAI GG(AG) TTI A(AG)C ATI GG.

Ces combinaisons d'amorces ont été conçues pour permettre d'amplifier des produits de l'ordre de 520 pb.

10 L'amplification a été réalisée dans 50 microlitres d'une solution contenant 5 microlitres d'ADNc, 2 mM dNTP, 100 pmol de chaque amorce dégénérée, 1,5 U de polymérase Taq (Boehringer Mannheim, Allemagne), 50 mM KCl, 2,5 mM MgCl₂, 10 mM Tris/HCl
15 pH8,3 et 0,01 de gélatine. Pour éviter l'évaporation, la surface du mélange a été couverte par 35 microlitres d'huile minérale (Sigma, France). La PCR a été réalisée à l'aide d'un thermocycler (Hybaid, Omnigene, USA) selon le protocole suivant : un cycle à 94°C pendant 90 s, 40
20 cycles à 94°C pendant 20 s, 50°C pendant 25 s et 72°C pendant 90 s, et un cycle à 72°C pendant 120 s.

Après la PCR, 5 microlitres du produit de la réaction ont été analysés sur gel d'agarose Seaplaque 2%, pour vérifier que la présence du fragment (Tebu).
25 S'il était présent, les 45 microlitres restant ont été soumis à électrophorèse et l'ADNc a été extrait du gel d'agarose à l'aide du kit QIARX II (Qiagen). L'ADNc extrait a été inséré dans le vecteur pMOSBlue qui a été utilisé pour infecter les cellules de E.coli MOSBlue
30 compétentes en utilisant le kit pMOSBlue T-vector selon le protocole du fournisseur (Amersham). Les bactéries infectées ont ensuite été cultivées sur un milieu sélectif (Xgal/IPTG).

Les clones recombinants ont été testés par
35 PCR directe sur colonie. En bref, chaque colonie blanche a été resuspendue dans 10 microlitres de tampon TE. La

PCR a été réalisée dans 10 microlitres d'une solution contenant 1 microlitre de suspension de colonie, 3 pmoles de chaque amorce universelle U19 et T7, 10 mM dNTP, 50 mM KCl et 2,5 mM MgCl₂ dans un tampon Tris HCl pH 8,3 avec 0,25 U de polymérase Taq. Le protocole pour la PCR était le suivant : un cycle à 94°C pendant 270 s, 30 cycles à 94°C pendant 30 s, 48°C pendant 30 s et 72°C pendant 50 s, et un cycle à 72°C pendant 120 s. Après la PCR, 10 microlitres du produit de la réaction ont été analysés sur un gel à 2% d'agarose. Les clones positifs ont été cultivés en milieu LB liquide contenant 0,1 mg / ml d'ampicilline.

5) Extraction et purification des fragments d'ADNc.

L'ADNc plasmidique a été extrait et purifié en utilisant le kit Wizard miniprep (Promega). Les échantillons ont été séquencés par Genome Express (Grenoble, France).

6) Analyse des séquences.

La comparaison des séquences de récepteur olfactif de marmotte de l'invention avec d'autres séquences disponibles dans GenBank / GenPept a été réalisée en utilisant le programme Blast sur le serveur NCBI. ClustalW a été utilisé pour construire les alignements multiples et réaliser l'analyse phylogénétique. Les domaines hydrophobes ont été délimités en utilisant un simple profil d'hydropathie, et la prédiction des domaines transmembranaires α -hélicoidaux en utilisant le serveur PHD. Enfin, la variabilité des 14 séquences de marmotte alignées, ainsi que leur hydropathie moyenne, ont été déterminées et traduites sous forme graphique à l'aide du programme Rav3. Les domaines transmembranaires ont été prédits avec le logiciel Top Pred II.

II - RÉSULTATS.1) Isolement de l'ARNm.

5 Un échantillon d'approximativement 2 g, contenant principalement de l'épithélium olfactif et le cartilage le supportant a été retiré de la tête congelée d'une marmotte. Cet échantillon a été utilisé pour la purification et les tests d'ARNm selon la description de la section Matériel et Méthodes. Au total, 1,95 microgramme d'ARNm ont été obtenus. Pour augmenter les chances de cloner des récepteurs olfactifs, la moitié de l'ARNm obtenu a été transcrit en présence de l'amorce oligo d(T) et l'autre moitié en présence de l'amorce random (R).

10

15

2) Amplification des séquences de récepteurs olfactifs.

L'amplification par PCR a été réalisée avec 150 ng d'ARNm en utilisant les trois jeux d'amorces spécifiques dégénérées (c-t, 4-1, 3-2) décrites précédemment dans Matériel et Méthodes. L'analyse de l'électrophorèse réalisée avec des aliquotes de 5 microlitres des produits de la PCR a révélé des bandes uniques de la taille attendue (Fig.1). Avec l'ADNc "T", une bande 520 pb a été obtenue avec les amorces 3-2 et une bande de 720 pb avec les amorces c-t. Avec l'ADNc "R", une bande de 720 pb a été obtenue en utilisant les amorces c-t. Aucune bande n'a été observée dans les trois autres pistes. Dans les PCR de contrôle, dans lesquelles une seule amorce était utilisée, aucune bande de la longueur attendue n'a été observée. L'électrophorèse a été répétée en utilisant les 45 microlitres restant de l'échantillon, et les fragments de 550 et 720 pb ont été extraits. Etant donnée la diversité des récepteurs olfactifs, il a été supposé que

20

25

30

35

la population d'ADNcdans une bande était hétérogène et il n'a donc pas été essayé de séquencer directement les fragments d'ADNc amplifiés par PCR. Ces fragments ont été clonés dans E. coli comme décrit précédemment.

5

3) Clonage.

Après insertion dans le vecteur p-Mosblue et l'infection de E. coli MOSBlue compétentes, 139 clones bactériens ont été obtenus au total, dont 58 à partir de la PCR réalisée à partir de l'ADNc "R" et les amorces c-t (clones R c-t), 31 à partir de la PCR réalisée à partir de l'ADNc "T" et les amorces c-t (clones T c-t) et 50 à partir de la PCR réalisée à partir de l'ADNc "T" et les amorces 3-2 (clones T 3-2). Pour confirmer la présence du fragment attendu, nous avons réalisé une autre PCR sur chacun des 139 clones en utilisant des amorces correspondant aux zones du vecteur situées sur chaque côté du fragment. L'électrophorèse sur gel d'agarose des produits de la PCR a montré que 5 clones R c-t, 10 clones T c-t et 22 clones T 3-2 possédaient des fragments de la taille attendue. Ces 37 clones positifs ont été cultivés à nouveau pour une production de masse.

10

15

20

4) Séquencage.

L'ADN plasmidique a été extrait, purifié et séquencé, comme décrit précédemment. Les séquences nucléotidiques ont été comparées avec celles trouvées dans les banques de données. Sur les 28 séquences présentant des scores élevés de similarité avec les récepteurs olfactifs, 14 étaient différentes et ininterrompues (AMOR 1 à 14) et pouvaient coder pour des récepteurs olfactifs. Les autres 14 séquences étaient identiques (n=8), inutilisables (n=3) ou incomplètes pour nos conditions expérimentales (116, 153, 159 acides aminés). Les 14 séquences utilisables avaient un seul cadre ouvert de lecture permettant leur traduction en

25

30

35

acides aminés. L'attribution de la séquence de lecture correcte a été confirmée par la similarité de ces traductions putatives avec les séquences en acides aminés d'autres récepteurs olfactifs disponibles dans Gen Bank / GenPept. Le pourcentage de résidus identiques dans les meilleurs alignements s'étendait entre 84% (entre AMOR4 et une séquence partielle de *Xenopus laevis* No d'accession #:1617233) et 46 % (entre AMOR5 et la séquence de *Rattus norvegicus* No d'accession#:1016362). 7 des 14 séquences de marmotte présentaient le meilleur alignement avec différents récepteurs de rat, 3 avec le même récepteur humain (accession : #AC002988), 3 avec la même séquence de chien (accession #:X89660) et un avec la séquence de Xénope citée précédemment. Le pourcentage moyen de résidus identiques était 64%. Sept (AMOR 1-7) des nouvelles séquences de marmotte ont été amplifiées à partir d'un couple d'amorces conçu à partir des domaines transmembranaires II et VII et sont longs de 234 à 237 résidus. Sept autres séquences (AMOR 8-14) ont été obtenues avec les amorces conçues à partir de la boucle intracellulaire 2 (i2) et le domaine transmembranaire VII et contiennent 176 résidus. Le pourcentage de résidus identiques entre ces 14 nouvelles séquences est compris entre 33 % (AMOR 4 / AMOR 8) et 79 % (AMOR 8 / AMOR 11).

5) Structure du domaine du récepteur olfactif putatif de marmotte.

L'homologie globale entre les 14 nouvelles séquences de marmotte et les séquences de récepteurs précédemment identifiées laisse peu de doutes quant à leur appartenance à la même superfamille des récepteurs à 7 domaines transmembranaires. Selon la localisation des amorces utilisées pour les amplifier, les séquences partielles AMOR 1-7 et AMOR 8-14 devraient présenter 6 ou 4 domaines transmembranaires respectivement. La

Figure en annexe 3 montre que le profil d'hydrophobicité de ces séquences est compatible avec une telle organisation. Afin de délimiter plus précisément les régions transmembranaires α -hélicoidales, l'alignement de la Figure 2 a aussi été soumise au serveur PHD. 5 régions transmembranaires ont été assignées sans ambiguïté dans les régions respectives (38-62), (86-103), (140-164), (186-203) et (216-232), qui correspondent aux domaines DIII, DIV, DV, DVI et DVII dans la figure 2.

Les inventeurs ont également chercher à situer les positions impliquées dans le site de fixation spécifique de l'odeur en appliquant une analyse précédemment introduite pour les molécules qui lient les antigènes. Ici, le raisonnement est que si ces récepteurs olfactifs sont supposés lier spécifiquement des molécules odorantes, les résidus qui constituent le site spécifique de liaison pourraient montrer plus de variabilité que ceux qui sont impliqués dans la structure core et dans la fonction de signallement.

La Figure 4 montre les profils de variabilité obtenus avec l'alignement de la Figure 2. Quatre pics de variabilité sont clairement visibles. Le profil moyen d'hydropathie (average hydropathy plot) montré en parallèle (Figures 2 et 4) indique qu'ils ne sont pas uniquement situés à l'intérieur de boucles hydrophiles comme escompté (position 210), mais aussi dans des régions hydrophobes (e.g. position 148). Le centre des segments les plus variables est situé aux positions 30, 100, 148 et 210, la cartographie respectivement à l'intérieur de la 1ère boucle extracytoplasmique E1, les 4ème et 5ème régions transmembranaires DIV et DV, et le milieu de la 3ème boucle extracytoplasmique E3. Nous proposons que les résidus à ces positions pourraient être impliqués dans le site de liaison de molécules odorantes inconnues

correspondant à ces récepteurs. Ces positions sont compatibles avec l'hypothèse selon laquelle les régions transmembranaires pourraient s'assembler en un calice ouvert vers l'extérieur pouvant recevoir une molécule odorante. Un tel modèle est aussi en accord avec le fait que beaucoup de molécules odorantes présentent un caractère hydrophobe.

6) Classification structurale des récepteurs olfactifs.

Nous avons tenté de classer les récepteurs clonés de marmotte par rapport aux séquences précédemment décrites pour d'autres espèces. La Figure 5 montre une classification structurale de 122 récepteurs olfactifs de la banque de données EMBL trouvés dans différentes espèces ainsi que les 14 séquences complètes et les 3 séquences incomplètes identifiées chez la marmotte dans le cadre de la présente invention. A l'exception des récepteurs de poisson, les récepteurs ne sont pas regroupés par espèces. Il y a 5 familles contenant un nombre varié de récepteurs. Les récepteurs olfactifs de marmotte ont été classés en sous-familles 1, 2 et 5. 12 séquences ont été rangées dans la sous famille 2.

Le plus fort pourcentage d'homologies interespèces (plus de 70 % de résidus identiques) entre récepteurs olfactifs a été observé dans 9 cas indiqués par une astérisque : entre le rat et la souris (jusqu'à 95 %) dans 5 cas, entre le rat et l'homme (80 %) dans un cas, entre le chien et l'homme (jusqu'à 85 %) dans deux cas, et entre le récepteur de marmotte et celui de rat dans un cas (73 %). L'homologie entre les récepteurs humains et de marmotte ne dépasse jamais 75 % de résidus identiques.

III - DISCUSSION.

Les récepteurs olfactifs comprennent une large famille multigénique. Leur étude demande une combinaison d'approches. Une stratégie de PCR inverse avec plusieurs amorces différentes a été mise en oeuvre dans le cadre de la présente invention. Cette approche a été couronnée de succès puisque 28 séquences putatives de récepteurs olfactifs dont 14 pouvaient permettre une analyse comparative ont été obtenues. Il est possible d'obtenir plus de séquences en changeant simplement les conditions de PCR. La famille de gènes clonés dans le cadre de la présente invention code pour des récepteurs olfactifs pour deux raisons. D'une part les profils d'hydropathie des séquences sont en accord avec les récepteurs de la superfamille des récepteurs à sept domaines transmembranaires. D'autre part, la comparaison avec les séquences de banques de données montre un fort degré de similarité avec les récepteurs olfactifs précédemment identifiés.

Les sites potentiels de reconnaissance des ligands sur les récepteurs olfactifs putatifs de marmotte ont été identifiés. Comme l'olfaction requiert la reconnaissance spécifique d'une grande variété de molécules odorantes, il a été postulé que le site de liaison du récepteur olfactif avec son ligand présenterait une plus grande variabilité entre résidus que les autres parties de la séquence responsables de la structure core et de la fonction de transduction. Il a été observé la plus forte variabilité à l'intérieur de deux domaines transmembranaires (DIV et DV) et à l'intérieur de deux boucles extracellulaires (E1 et E3). Il a donc été conclu que ces régions pourraient être impliquées dans la reconnaissance du ligand.

La présence d'un site de liaison profond dans le calice transmembranaire n'est pas un caractère spécifique des récepteurs olfactifs récepteur mais est

commun parmi les récepteurs à 7 domaines transmembranaires des amines biogènes.

Le site principal d'interaction entre les récepteurs à 7 domaines transmembranaires et la protéine G apparentée est la troisième boucle intracellulaire. Pour les séquences présentées ici, le segment le plus conservé est situé entre les positions 180 à 193, c'est à dire la fin de cette boucle et le début du 6ème domaine transmembranaire.

Les résultats obtenus indique une analogie remarquable entre le récepteur olfactif de marmotte et le récepteur olfactif de rat. La longueur (18 résidus) de la 3ème boucle intracellulaire (i3) était courte. La séquence consensus IVSSI (ou une séquence proche) était à l'extrémité N-terminale de la 3ème boucle intracellulaire dans 75 % des clones de l'invention. La troisième boucle intracellulaire est riche en résidus Serine et peut donc constituer des sites de phosphorylation pour GRK. Les récepteurs à 7 domaines transmembranaires sont classés en plusieurs groupes. Les récepteurs olfactifs sont supposés appartenir au groupe I, qui est caractérisé par la présence d'une séquence DRY strictement conservée du côté N-terminal de i2. La séquence DRY est présente dans 4 des clones de l'invention mais est remplacée par une séquence DRF dans les 10 restant.

La reconnaissance des mêmes odeurs par des espèces différentes soulève une intéressante question. On peut s'attendre à ce que ces espèces aient des récepteurs orthologues. En utilisant le logiciel clustalW (figure 5), les inventeurs ont cherché à déterminer si certains des récepteurs olfactifs de marmotte étaient bonafide orthologues de récepteurs olfactifs d'autres espèces, en particulier d'autres rongeurs. Pour les récepteurs couplés aux protéines G, les pourcentages d'identité entre les récepteurs

orthologues de différentes espèces allaient de 68 %
(pour le récepteur CSN entre le chien et l'homme) à 98 %
(pour le récepteur cannabinoïde du rat et de l'homme).
Des récepteurs olfactifs avec des pourcentages de
5 similitude de cet ordre ont été observés entre rat et
souris, rat et homme, et chien et homme. Un seul
récepteur olfactif de marmotte présentait un pourcentage
de similitude de cet ordre avec un récepteur de rat
(AMOR14 73%). En général, nous avons trouvé peu
10 d'homologies fortes. Cette découverte pourrait indiquer
que, soit le nombre de récepteurs olfactifs était trop
petit pour permettre l'identification de vrais
récepteurs orthologues, soit le pourcentage de
similarité entre récepteurs olfactifs orthologues peut
15 devenir inférieur à 68 %.

Une autre alternative serait que les animaux
sauvages expriment de récepteurs pour davantage d'odeurs
que les animaux de laboratoire. La marmotte des Alpes
(*Marmota marmota*) a été choisie comme modèle dans cette
20 étude à partir de l'hypothèse selon laquelle, étant
donnée l'importance de l'olfaction pour survivre dans la
nature, l'olfaction serait fortement développée. La
marmotte des Alpes marque son territoire avec des
sécrétions produites par les glandes jugales. De plus,
25 pour cet animal, le sens olfactif est de la plus grande
importance parce que cette espèce possède un fort niveau
de sociabilité : il vit dans des groupes familiaux
formés par une paire d'adultes résidents reproductifs et
leur progéniture de plusieurs portées successives qui
30 restent dans le groupe natal jusqu'à l'âge de 2 ans ou
plus. Chaque marmotte a une combinaison différente de
molécules odorantes que les membres du même groupe ou
d'un groupe différent peuvent flairer.

Contrairement à d'autres systèmes
35 sensoriels, le système olfactif requiert une myriade de
récepteurs différents. Comme les mammifères sont

généralement supposés avoir environ un millier de gènes, les clones identifiés dans cette étude représentent probablement seulement une partie de la famille des récepteurs olfactifs de marmotte. En plus de la contribution au nombre de récepteurs identifiés, nos résultats supportent aussi l'existence de récepteurs orthologues entre espèces et la notion selon laquelle la variabilité locale observée dans certains des domaines transmembranaires pourrait être capitale pour la spécificité d'un récepteur. Comment même un millier de récepteurs pourrait être capable de distinguer parmi la dizaine de milliers d'odeurs trouvées dans la nature n'est pas encore clair. La confirmation finale de la nature et de la spécificité olfactive de ces récepteurs ne sera pas possible tant que la séquence entière n'a pas été obtenue et la liaison spécifique avec une ou plusieurs molécules odorantes démontrée.

RÉFÉRENCES BIBLIOGRAPHIQUES

(1) Axel R. (1995). The molecular logic of smell. *Scientific American* October 154-159.

5

(2) Bel M.C., Porteret C. and Coulon J. (1995). Scent deposition by cheek rubbing in the Alpine marmot (*Marmota marmota*) in the French Alps. *Can. J. Zool.*, **73**, 2065-2071.

10

(3) Buck L. and Axel R. (1991). A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell*, **65**, 175-187.

15

(4) Cornell B.A., Braach-Maksvytis, V.L.B., King L.G., Osman P.D.J., Raguse B., Wieczorek L. and Pace R.J. (1997) A biosensor that uses ion-channel switches. *Nature* 387, 580-583.

20

(5) Kinoshita T. (1995) Biomembrane mimetic systems. *Prog. Polym. Sci.*, **20**, 527-583.

(6) Mielle P. (1998). Une technique de pointe au service du contrôle de la qualité aromatique. *Biofuture*, **174**, cahier n°99

25

(7) Raming K., Krieger J., Strotmann J., Boekhoff I., Kubick S., Baumstark C., and Breer H. (1993). Cloning and expression of odorant receptors. *Nature*, **361**, 353-356.

30

(8) Wu T.T., and Kabat E.A. (1970). An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J. Exp. Med.*, **132**, 211-250.

35

REVENDICATIONS

1) Un récepteur olfactif purifié de marmotte.

5

2) Un récepteur olfactif constitué par ou comprenant une séquence en acides aminés choisie parmi celles représentées dans la liste de séquences en annexe sous les numéro SEQ ID No:1 à SEQ ID No:23, ou un dérivé fonctionnellement équivalent de celles-ci.

10

3) Un récepteur selon la revendication 2 constitué par ou comprenant une séquence en acides aminés présentant environ 75% et de préférence au moins 95% d'homologie avec une séquence en acides aminés choisie parmi celles représentées dans la liste de séquences en annexe sous les numéro SEQ ID No:1 à SEQ ID No:23.

15

4) Un récepteur selon l'une des revendications 2 ou 3, constitué par ou comprenant une séquence en acides aminés choisie parmi celles représentées dans la liste de séquences en annexe sous les numéro SEQ ID No:1 à SEQ ID No:23 dont une ou plusieurs des régions très hétérogènes est modifiée.

20

25

5) Un anticorps poly ou monoclonal dirigé contre au moins un récepteur selon l'une quelconque des revendications 1 à 4 ou un dérivé ou un fragment de ceux-ci.

30

6) Une molécule d'acide nucléique comprenant ou constituée par une séquence nucléique codant pour un récepteur selon l'une quelconque des revendications 1 à 4.

35

7) Une molécule d'acide nucléique selon la revendication 6, comprenant ou constituée par une séquence choisie parmi celles représentées dans la liste de séquences en annexe sous les numéros SEQ ID No:24 à
5 SEQ ID No:47.

8) Un vecteur comprenant au moins une molécule d'acide nucléique selon l'une des revendications 6 ou 7, avantageusement associée à des
10 séquences de contrôle.

9) Procédé de production d'un récepteur selon l'une quelconque des revendications 1 à 4, caractérisé en ce qu'il consiste :

- 15 - à transférer une molécule d'acide nucléique selon l'une des revendications 6 ou 7 ou un vecteur selon la revendication 8 dans un hôte,
- à cultiver ledit hôte cellulaire dans des conditions permettant la production de la protéine
20 constituant le récepteur,
- à isoler, par tous moyens appropriés les dites protéines.

10) Procédé d'expression d'un récepteur selon l'une quelconque des revendications 1 à 4 chez un hôte, caractérisé en ce qu'il consiste :

- 25 - à transférer une molécule d'acide nucléique selon l'une des revendications 6 ou 7 ou un vecteur selon la revendication 8 dans un hôte,
- 30 - à cultiver ledit hôte dans des conditions permettant l'expression desdites récepteurs à la surface de l'hôte.

11) Un hôte transformé par une molécule d'acide nucléique selon l'une des revendications 6 ou 7
35 ou par un vecteur selon la revendication 8.

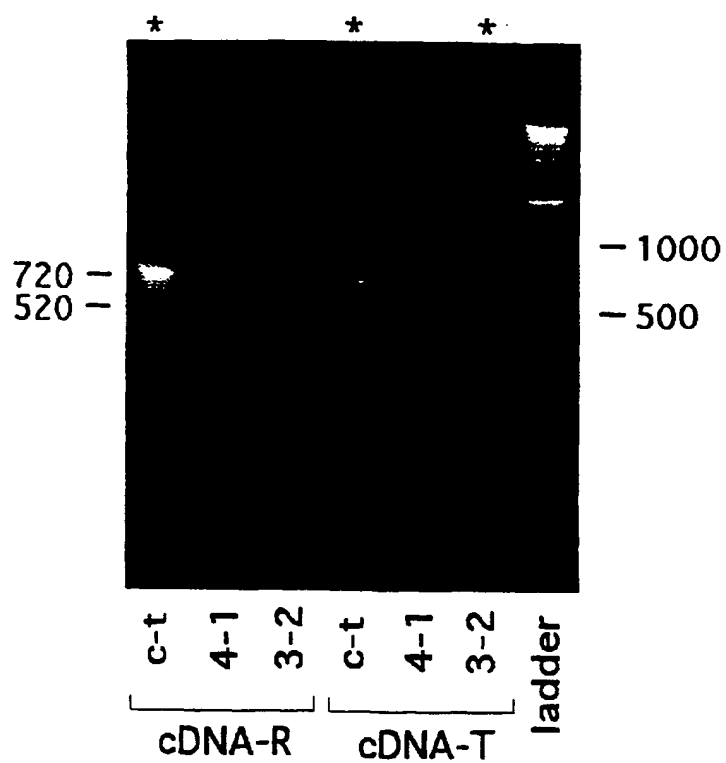
12) Procédé de criblage de composés susceptibles de constituer des ligands d'un récepteur selon l'une quelconque des revendications 1 à 4, caractérisé en ce qu'il consiste à mettre en contact un
5 composé et un ou plusieurs desdits récepteurs puis à mesurer par tout moyen approprié l'affinité entre ledit composé et ledit récepteur.

13) Une membrane sur laquelle est fixée un
10 ou plusieurs récepteurs selon l'une quelconque des revendications 1 à 4 utile pour la mise en oeuvre d'un procédé selon la revendication 12.

14) Un composé constituant un ligand d'un
15 récepteur olfactif, identifié et sélectionné par le procédé selon la revendication 13.

15) Utilisation d'un récepteur selon l'une
20 quelconque des revendications 1 à 4, d'un hôte selon la revendication 11 ou d'une membrane selon la revendication 13, pour la détection d'arômes, le contrôle de qualité, l'analyse d'échantillons, l'analyse ou la comparaison de parfums, la détection de substances toxiques, ou le piégeage d'odeurs.

Fig. 1



2 / 14

Fig. 2

	DII	E1	DI II
1	-----23-----43-----		
AMOR1	PMYFLGNLSFLDLSTSSIPQLLHNLGRDKTISYVGCVVQLFLFLGLGVECLLLA		
AMOR2	PMYFLGNLSFVEVCLTSTTVPKILVNTQTLTKDLSYRGCLTQVYFLMVFAGMDNFLLT		
AMOR3	PMYFLGNLSFLEVWYTTAAVPAKALAILLGRSQSISFISCLLQMYLVFSLGCTEYFLLV		
AMOR4	PMYFLGNLSFIDVCHSTVTPKMLRDTWSEKLLISFDACVTQMFFLHLFACTEIFLLT		
AMOR5	PMYFLGNLSLLEIGYTCVTPKMLQSLVSEARGISREGCATQMFFFTLFAISECCLLA		
AMOR6	PMYFLGNLSFLEILYTTSTVVPKMLEGFLQVA-AISVTGCLTQFFIFGSLATAECFLLA		
AMOR7	PRYFLGNLSLADIGISTTTPQMVVNIQKRRTISYAGCLTQVCVLI FAGSENFLLA		
AMOR8	-----Q-----		
AMOR9	-----Q-----		
AMOR10	-----Q-----		
AMOR11	-----Q-----		
AMOR12	-----Q-----		
AMOR13	-----Q-----		
AMOR14	-----Q-----		
	*****	* .	** . * . *

Fig. 2 -suite-

	i2	DIV	E2
AMOR1	--64-----83-----101-----		
AMOR2		VMA YDRFVA VCKPLHYTVIMSSRLCLGLSVAWGCGMANSLVSPVTLQLPRCGHNKVD	
AMOR3		VMA FDRFVA ICYPLNYTVIMNPRCLVLLVLLSWLIMFWVSLHILLKRLTFSSGTAVP	
AMOR4		AMAYDRYVAICFPPLHYTTIMSLKCLSLVLSWVLTMLHALLHTLLVVRLSFCSDNVIP	
AMOR5		VMA YDRYVAICKPLQYMTVMNWKVCVLLAVALWAGGTIHSISLSTIKLPYCGPDEID	
AMOR6		AMAFDRYMAICSPHYATRM SRGVCAHLAVVSWTVGCMVGLGQTN YIFSLDFCGPCEID	
AMOR7		VMA YDRFLAICYPLRYPLLMGPRWCMGLVVTAWLSGFMVDELVVVLM AQLRFCGSNRID	
AMOR8		AMAYDRYAAICHPRLRYTAIMNPHLCVLLVMISLSISTVDALLHSLMLLRLSFCTDLEIP	
AMOR9		ALAYDRFVAICHPHLHYLVIMSPRHCGFLT VLSFLLSLLDSQLHSFMTLNITSFKDV EIS	
AMOR10		ALAYDRFVAICYPLHYMVMINSRRCGLLILVSWIMSAHSL LQGLMMLRLSFCTDLEIS	
AMOR11		ALAYDRFVAICHPHLHYPRIMSQNL CFLLVVVSWVLSANALLHTLL LARLSFLRGITLP	
AMOR12		ALAYDRFLAICYPLHYTVIMNPRLCGFSILVSFLLSLLDSQLHNL MILQITSFKDV EIS	
AMOR13		ALAYDRFVAICHPHLHYPTIMNPRFCGFLVLSFVLSLLESQ LHNLI ALQFTTFKDDVKIA	
AMOR14		ALAYDRFGAIRFP LHNNTTIMSPKLGLEFLVLSWVLTMTFHAMLHTLL MARLCFAENMI P	
	*** ** ** *	ALAYDRFLAICHPHLHYTAIMNPRLCGLLVLCWILSVLHALLQSLMVLRLSFCDIEIP	

Fig. 2 -suite-

	E2	DV	i3
AMOR1	-----138-----	-----163-----	-----
AMOR2	HFCELPALIRMACVNTVAIEGTVFLAVGIVLSPLVFILVSYGHI VRAVFRIQSSSGR		
AMOR3	HFFCELSQLLKATSSDTLVNIIILYVVTALLGIFPATGILYSYSQIVSSLLRMSSSVGK		
AMOR4	HFSCEISALLKLACSNTHVNELVIFITGGLVIVTPFLLILGSYVQIFSSILKVPSARGI		
AMOR5	NFFCDVPQVIKLA CTDTHII EILIVSNSGLISVVCVFLVVSYAVILVSLRQQIS -EGR		
AMOR6	HFFCDLPPILALACGDTSHNEAAVFVVAI LCISPPFLIVASYGRILA AVLVMPSPEGR		
AMOR7	HFYCDFMPLVVLACSDPRVAQVTTFVLSVFLTVPFGLILTSYARI VTVLRVPAGASR		
AMOR8	HFFCELDQVITLACSDTLINNLLIYVTAGIFAGVPLSGIIFSYLHIVSSVLRMPSPGGV		
AMOR9	NFFCDPSQLLNLSCSNTFSDNIVKYFLGAFYGLFPI SGILFSYKIISSILRIPSLGGK		
AMOR10	HFFCELNHLVHLACSDTFLNEVVIYFAAVLLAGGPLAGILYSYCKIVSSIHAISSAQGK		
AMOR11	HFFCDLSALLKLSSSDTTINQLAILTAGSAVVTLPFMCILVSYGHIGATILRRPSLKG I		
AMOR12	SFFCDPSQLLNLSCSDNYSINTGKYVLFALYSFFPI SGILFSYKIISSILRIPSSGGK		
AMOR13	NFFCDPSQVLSLSCSGTFINIIVMYFVGALFGVFPISGILFSYKIVSTILRIPSSGGK		
AMOR14	HFFCDMSALLKLSCSNTHVNELVIFITAGLILLIPLVLILLSYGHIVSSILKVPSARGI		
	HFFCELNQVVQLACFDNLLNDIVMNFALVLLATCPLAGILYSYSKIVSSIRAISSAQGK		
	. * .	. . * * *	

Fig. 3 -suite-

	DVI	E3	DVII
AMOR1	181-----204-----219-----236aa		
AMOR2	HRIFNTCGSHLTVVSLFYGNIIYMYMQPGSRSSQDQKFLTLFYNIIVTPLLNPFIYSLRN		
AMOR3	SKAFSTCGSHLCVVSLFYGTGLGVHLSSAMNHP SQGNMIASVMLHC GHPMLNPIIYTLRN		
AMOR4	HKAFSTCGSHLSVVSLFYGTII GLYLCP SANNSTVKD TVVALMYTVVTPMLNPF IYTLRN		
AMOR5	RKALSTCA AHLTVVTLFLGHCIF IYSRPSTSLPE--DKVVS VFFTA VTPLLNPF IYSLRN		
AMOR6	RKALSTCSSHLLVVTLFYGS GSVTYLRPKASHSPGMDKLLALFYTVVTSMLNPIIYTLRN		
AMOR7	RKAFSTCSSHLAVVSTFYGTLMVLIVIPSAVHSQLLSKVFALLYTVVTFIENPIIYSFRN		
AMOR8	YKAFSTCGSHLSVVCLFYGTIFGVYISSAVTDSQRKGAVASVMYSVVPQMLNPIIYTLRN		
AMOR9	YKAFSTCGSHLAVVCLFLVTASTVYLGSAVSHSPRNDVVASLMTYTVVTPMLNPFICS---		
AMOR10	YKAFSTCASHLSVVSLFYCTSPGVYLSAVTQNSHSTATA SVMYSVVT PMLNPFICS---		
AMOR11	CKALSTCGSHLSVVSVYGA VIALYIVPSSNSTNDKDIAVSVLYTLVIPMLNPFICS---		
AMOR12	YKAFSTCGSHLAVFCLFLGTGTAVYFGSAVSHSPRENVSVMYTVVTPMLNPFICS---		
AMOR13	YKAFSTCGSHLSVVCLFYGTGFGVYLGSAVSHSRKSAVASVMYTVVTPMLNPFICS---		
AMOR14	HKTFSTCGSHLSVVSLFYGTVIGLYLCP SANNSTVKD TVMALMYTVVTPMLNPFICS---		
	YKAFSTCASHLSVVSLFYCTGLGVYLSAVSHSRSSATASVMYTVVTPMLNPFICS---		
	. ** ** ** * ** *	.. ** *

Fig. 3

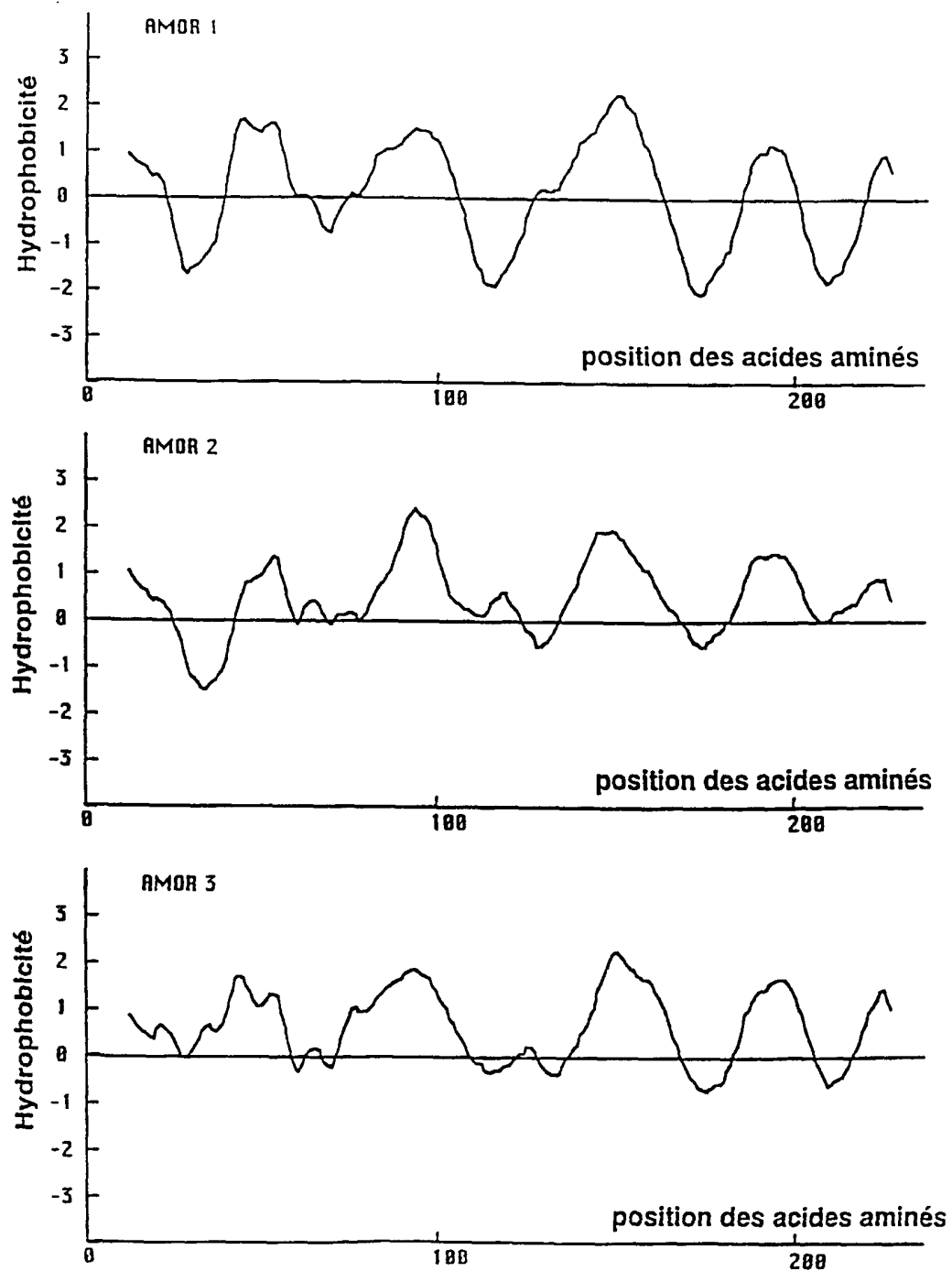


Fig. 3 -suite-

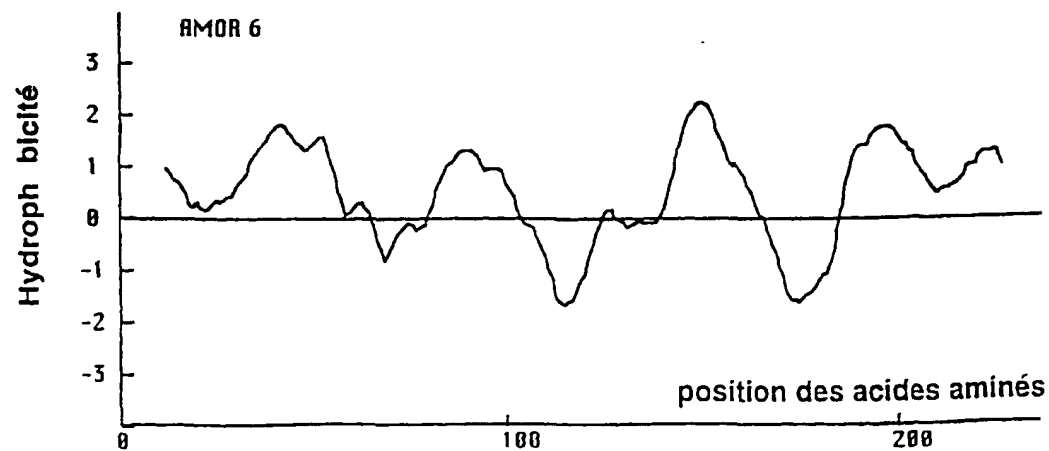
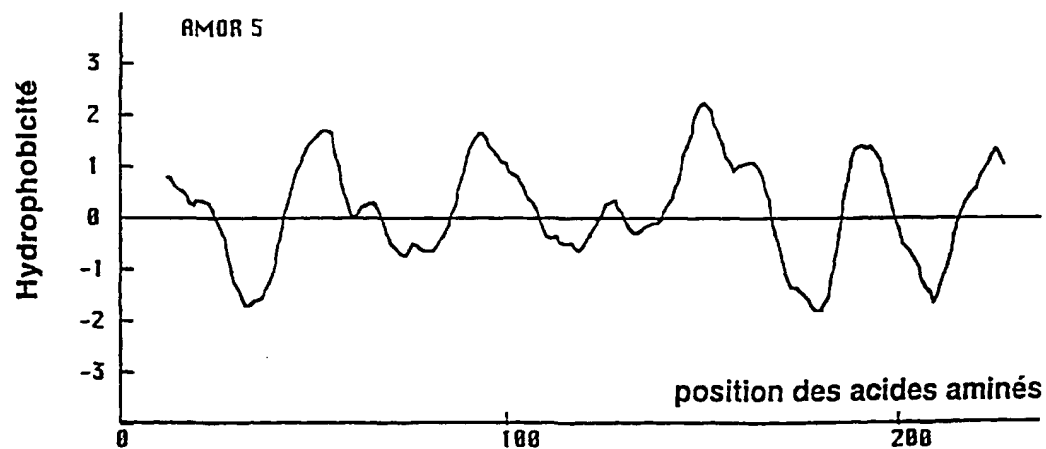
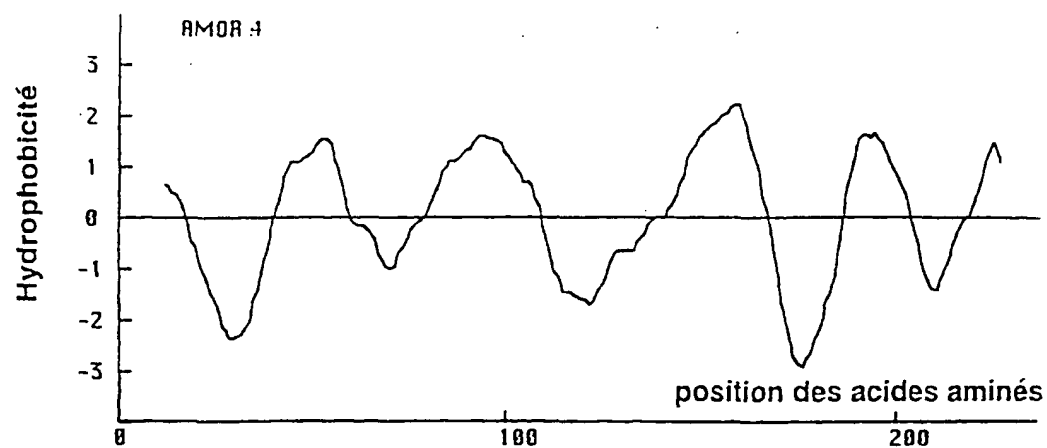


Fig. 3 -suite-

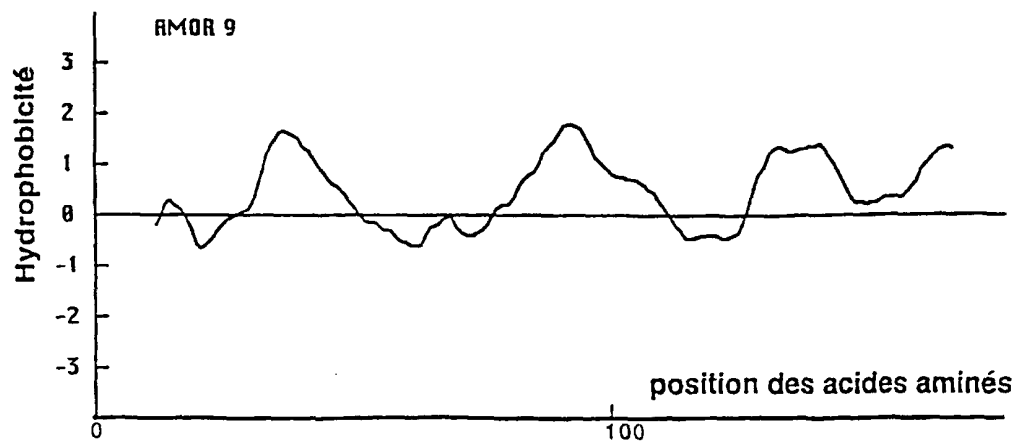
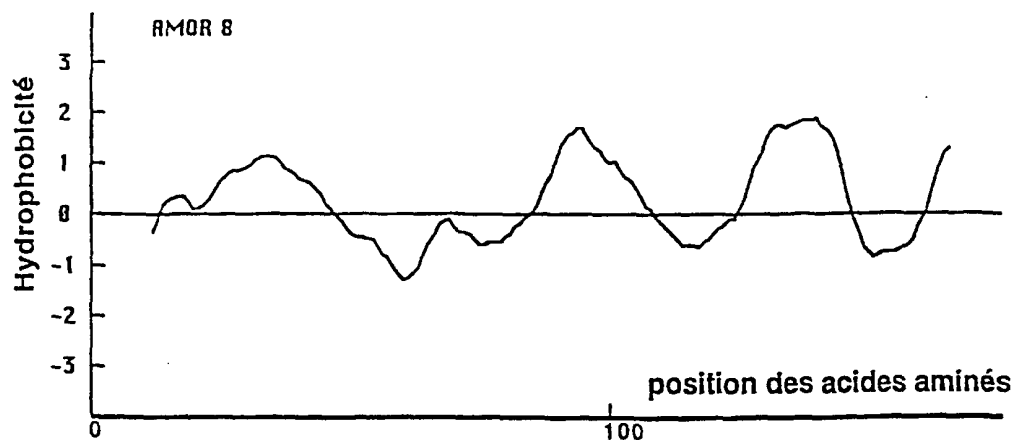
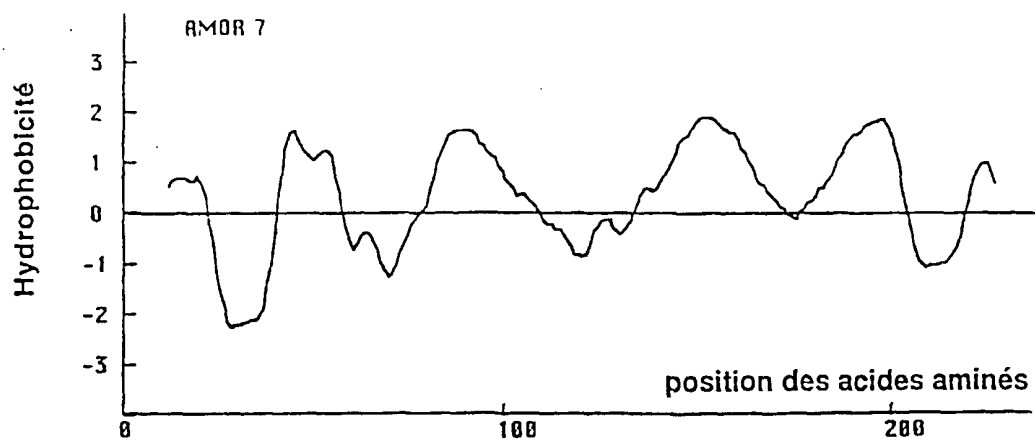


Fig. 3 -suite-

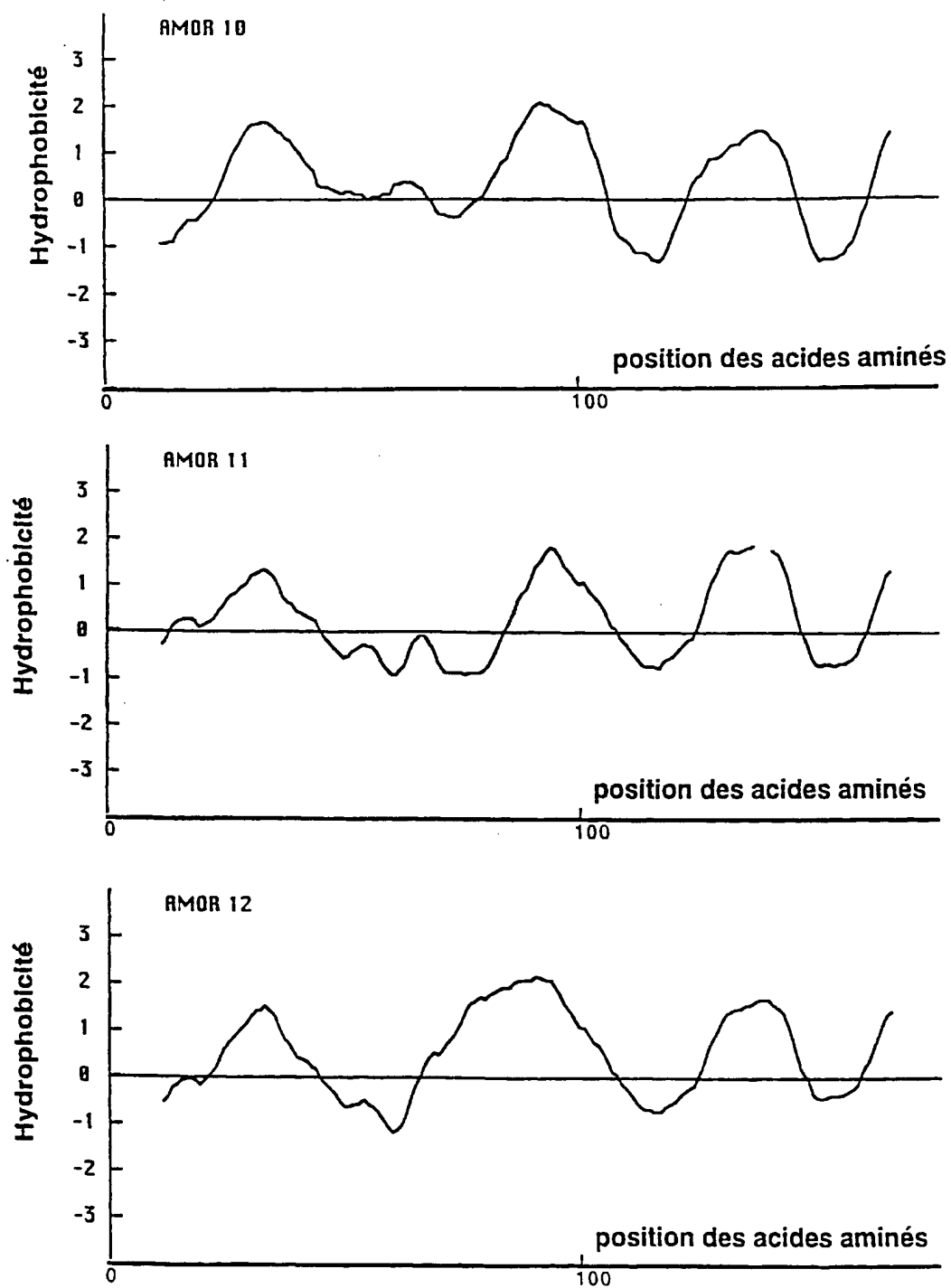


Fig. 3 -suite-

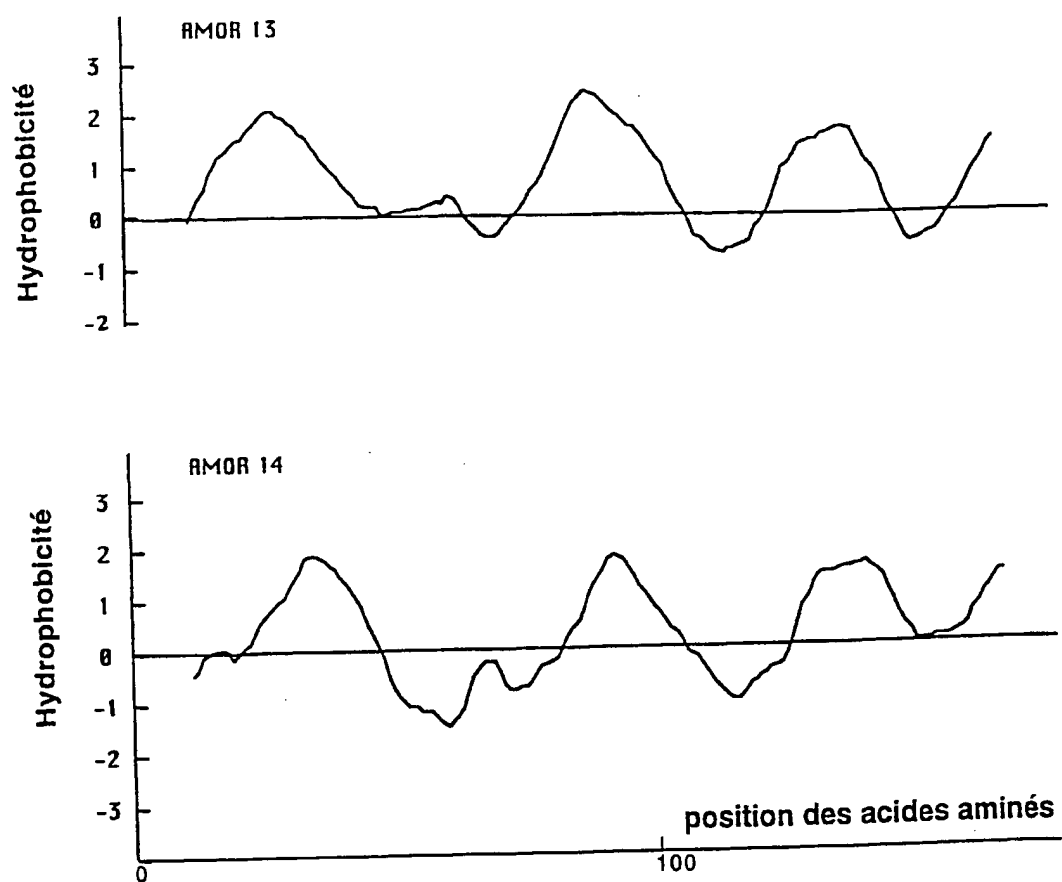
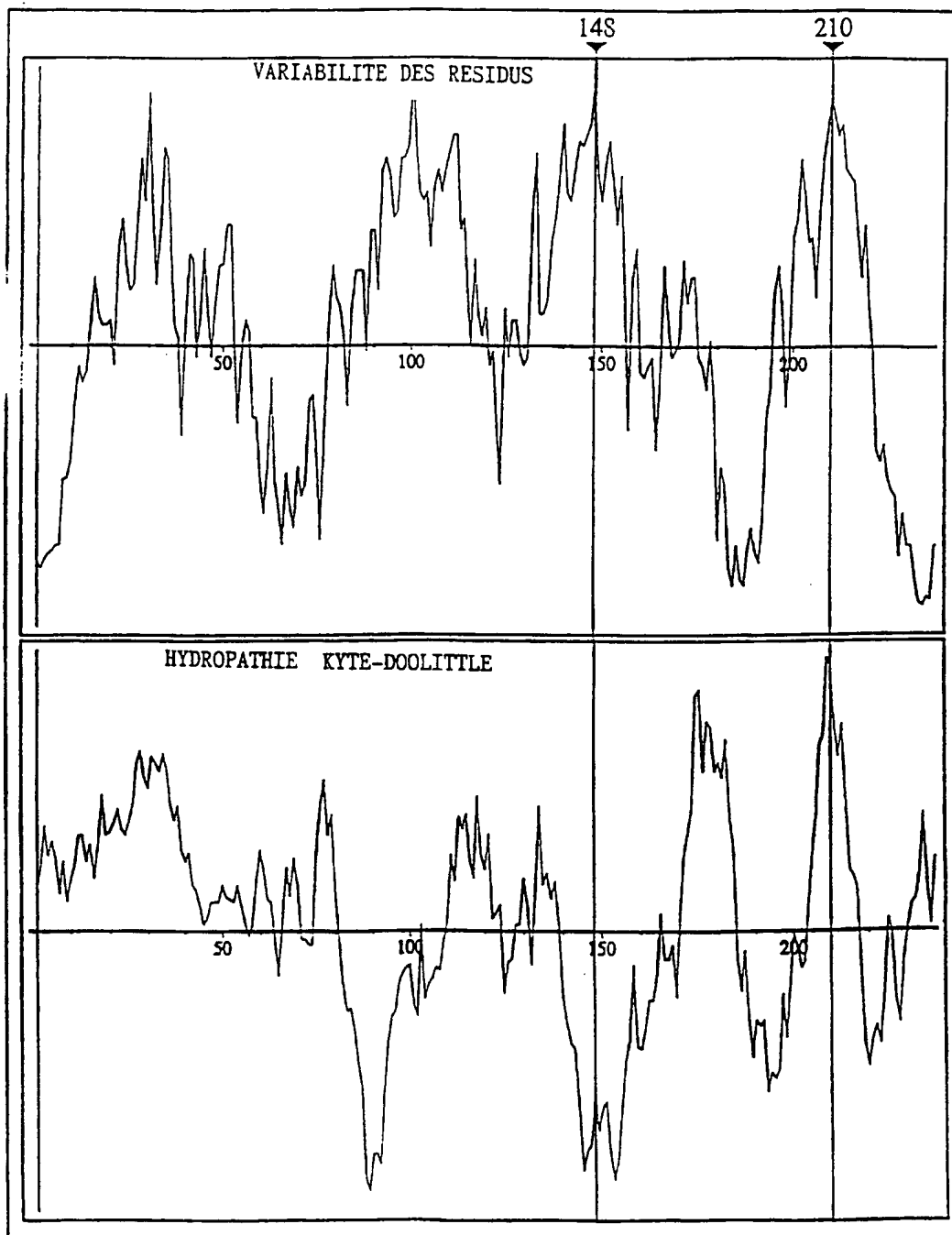


Fig. 4



13 / 14

Fig. 5 -suite-

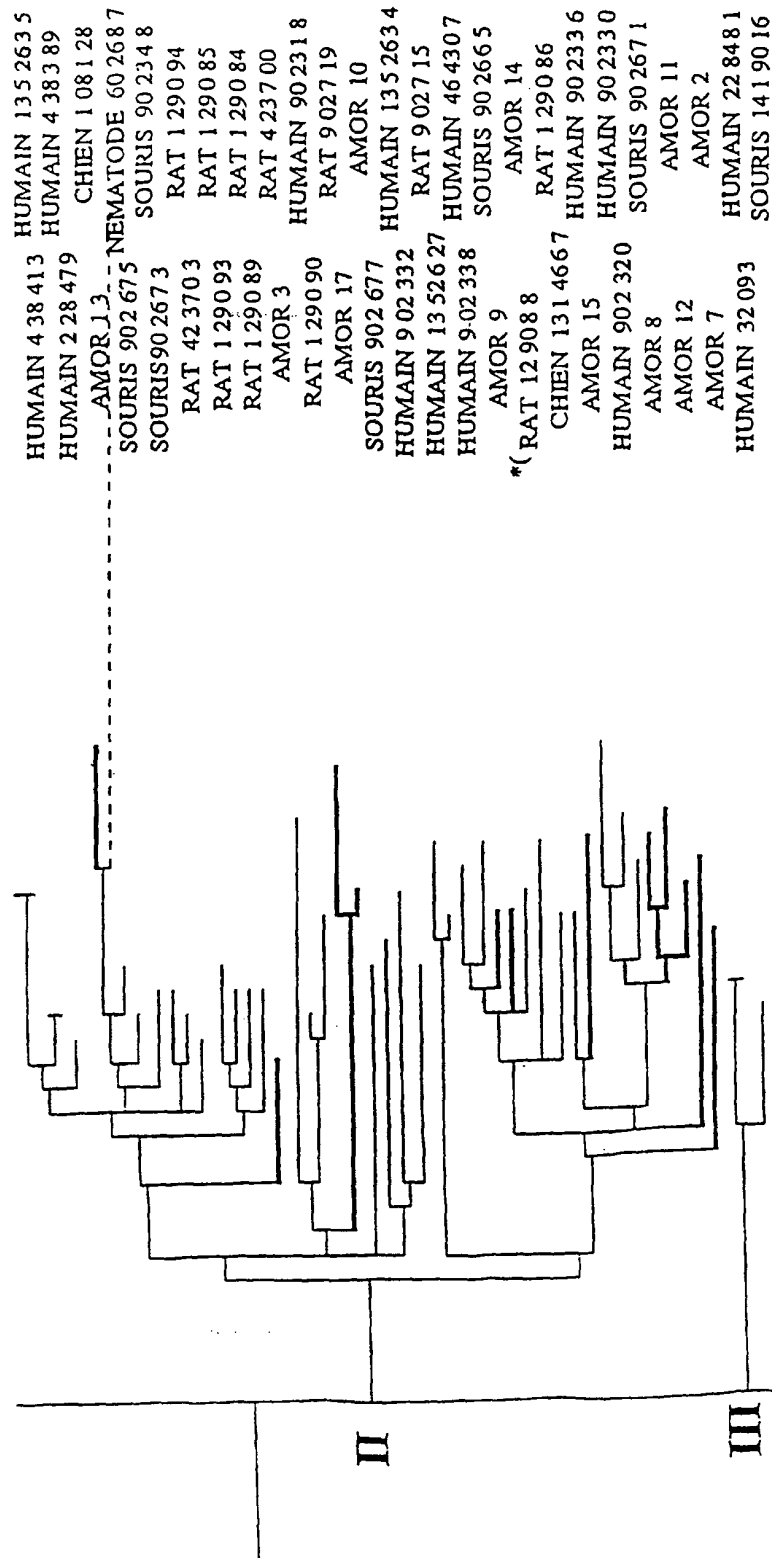
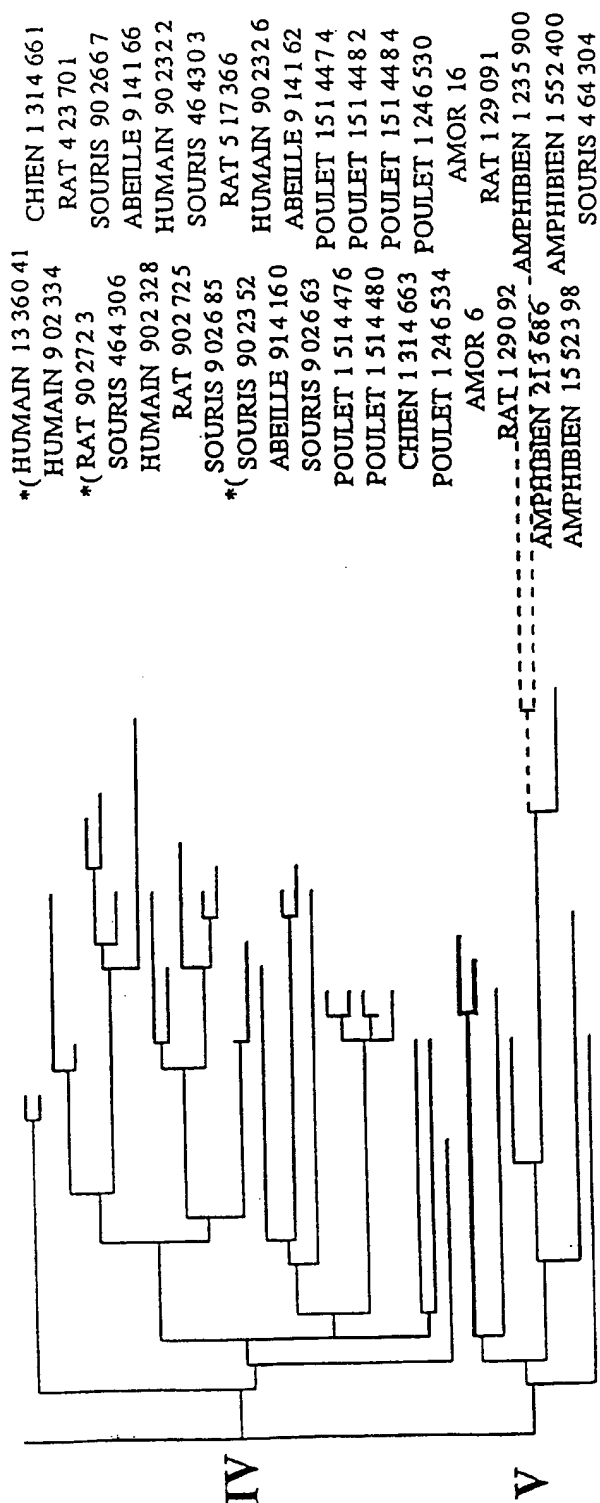


Fig. 5 -suite-



LISTE DE SÉQUENCES

(1) INFORMATION GÉNÉRALES:

(iii) NOMBRE DE SEQUENCES: 46

(2) INFORMATION POUR LA SEQ ID NO:1 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 237

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR1

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:1 :

Pro Met Tyr Leu Phe Leu Gly Asn Leu Ser Phe Leu Asp Leu Ser Phe
1 5 10 15

Thr Thr Ser Ser Ile Pro Gln Leu Leu His Asn Leu Ser Gly Arg Asp
20 25 30

Lys Thr Ile Ser Tyr Val Gly Cys Val Val Gln Leu Phe Leu Phe Leu
35 40 45

Gly Leu Gly Gly Val Glu Cys Leu Leu Leu Ala Val Met Ala Tyr Asp
50 55 60

Arg Phe Val Ala Val Cys Lys Pro Leu His Tyr Thr Val Ile Met Ser
65 70 75 80

Ser Arg Leu Cys Leu Gly Leu Val Ser Val Ala Trp Gly Cys Gly Met
85 90 95

Ala Asn Ser Leu Val Met Ser Pro Val Thr Leu Gln Leu Pro Arg Cys
100 105 110

Gly His Asn Lys Val Asp His Phe Leu Cys Glu Met Pro Ala Leu Ile
115 120 125

Arg Met Ala Cys Val Asn Thr Val Ala Ile Glu Gly Thr Val Phe Val
130 135 140

Leu Ala Val Gly Ile Val Leu Ser Pro Leu Val Phe Ile Leu Val Ser
145 150 155 160

Tyr Gly His Ile Val Arg Ala Val Phe Arg Ile Gln Ser Ser Ser Gly
165 170 175

Arg His Arg Ile Phe Asn Thr Cys Gly Ser His Leu Thr Val Val Ser
180 185 190

Leu Phe Tyr Gly Asn Ile Ile Tyr Met Tyr Met Gln Pro Gly Ser Arg

195	200	205
Ser Ser Gln Asp Gln Gly Lys Phe Leu Thr Leu Phe Tyr Asn Ile Val		
210	215	220
Thr Pro Leu Leu Asn Pro Phe Ile Tyr Ser Leu Arg Asn		
225	230	235 237

(3) INFORMATION POUR LA SEQ ID NO:2 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 237

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR2

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:2 :

Pro Met Tyr Leu Phe Leu Gly Asn Leu Ser Phe Val Glu Val Cys Leu		
1	5	10 15
Thr Ser Thr Thr Val Pro Lys Ile Leu Val Asn Thr Gln Thr Leu Ser		
	20	25 30
Lys Asp Ile Ser Tyr Arg Gly Cys Leu Thr Gln Val Tyr Phe Leu Met		
	35	40 45
Val Phe Ala Gly Met Asp Asn Phe Leu Leu Thr Val Met Ala Phe Asp		
	50	55 60
Arg Phe Val Ala Ile Cys Tyr Pro Leu Asn Tyr Thr Val Ile Met Asn		
	65	70 75 80
Pro Arg Leu Cys Val Leu Leu Val Leu Leu Ser Trp Leu Ile Met Phe		
	85	90 95
Trp Val Ser Leu Leu His Ile Leu Leu Leu Lys Arg Leu Thr Phe Ser		
	100	105 110
Ser Gly Thr Ala Val Pro His Phe Phe Cys Glu Leu Ser Gln Leu Leu		
	115	120 125
Lys Ala Thr Ser Ser Asp Thr Leu Val Asn Ile Ile Leu Leu Tyr Val		
	130	135 140
Val Thr Ala Leu Leu Gly Ile Phe Pro Ala Thr Gly Ile Leu Tyr Ser		
	145	150 155 160
Tyr Ser Gln Ile Val Ser Ser Leu Leu Arg Met Ser Ser Ser Val Gly		
	165	170 175
Lys Ser Lys Ala Phe Ser Thr Cys Gly Ser His Leu Cys Val Val Ser		
	180	185 190
Leu Phe Tyr Gly Thr Gly Leu Gly Val His Leu Ser Ser Ala Met Asn		
	195	200 205

His Pro Ser Gln Gly Asn Met Ile Ala Ser Val Met Leu His Cys Gly
 210 215 220

His Pro Met Leu Asn Pro Ile Ile Tyr Thr Leu Arg Asn
 225 230 235 237

(4) INFORMATION POUR LA SEQ ID NO:3 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 237

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR3

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:3 :

Pro Met Tyr Leu Phe Leu Gly Asn Leu Ser Phe Leu Glu Val Trp Tyr
 1 5 10 15

Thr Thr Ala Ala Val Pro Lys Ala Leu Ala Ile Leu Leu Gly Arg Ser
 20 25 30

Gln Ser Ile Ser Phe Ile Ser Cys Leu Leu Gln Met Tyr Leu Val Phe
 35 40 45

Ser Leu Gly Cys Thr Glu Tyr Phe Leu Leu Val Ala Met Ala Tyr Asp
 50 55 60

Arg Tyr Val Ala Ile Cys Phe Pro Leu His Tyr Thr Thr Ile Met Ser
 65 70 75 80

Leu Lys Leu Cys Leu Ser Leu Val Val Leu Ser Trp Val Leu Thr Met
 85 90 95

Leu His Ala Leu Leu His Thr Leu Leu Val Val Arg Leu Ser Phe Cys
 100 105 110

Ser Asp Asn Val Ile Pro His Phe Ser Cys Glu Ile Ser Ala Leu Leu
 115 120 125

Lys Leu Ala Cys Ser Asn Thr His Val Asn Glu Leu Val Ile Phe Ile
 130 135 140

Thr Gly Gly Leu Val Ile Val Thr Pro Phe Leu Leu Ile Leu Gly Ser
 145 150 155 160

Tyr Val Gln Ile Phe Ser Ser Ile Leu Lys Val Pro Ser Ala Arg Gly
 165 170 175

Ile His Lys Ala Phe Ser Thr Cys Gly Ser His Leu Ser Val Val Ser
 180 185 190

Leu Phe Tyr Gly Thr Ile Ile Gly Leu Tyr Leu Cys Pro Ser Ala Asn
 195 200 205

Asn Ser Thr Val Lys Asp Thr Val Val Ala Leu Met Tyr Thr Val Val
210 215 220

Thr Pro Met Leu Asn Pro Phe Ile Tyr Thr Leu Arg Asn
225 230 235 237

(5) INFORMATION POUR LA SEQ ID NO:4 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 234

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR4

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:4 :

Pro Met Tyr Leu Phe Leu Gly Asn Leu Ser Phe Ile Asp Val Cys His
1 5 10 15

Ser Thr Val Thr Val Pro Lys Met Leu Arg Asp Thr Trp Ser Glu Glu
20 25 30

Lys Leu Ile Ser Phe Asp Ala Cys Val Thr Gln Met Phe Phe Leu His
35 40 45

Leu Phe Ala Cys Thr Glu Ile Phe Leu Leu Thr Val Met Ala Tyr Asp
50 55 60

Arg Tyr Val Ala Ile Cys Lys Pro Leu Gln Tyr Met Thr Val Met Asn
65 70 75 80

Trp Lys Val Cys Val Leu Leu Ala Val Ala Leu Trp Ala Gly Gly Thr
85 90 95

Ile His Ser Ile Ser Leu Thr Ser Leu Thr Ile Lys Leu Pro Tyr Cys
100 105 110

Gly Pro Asp Glu Ile Asp Asn Phe Phe Cys Asp Val Pro Gln Val Ile
115 120 125

Lys Leu Ala Cys Thr Asp Thr His Ile Ile Glu Ile Leu Ile Val Ser
130 135 140

Asn Ser Gly Leu Ile Ser Val Val Cys Phe Val Val Leu Val Val Ser
145 150 155 160

Tyr Ala Val Ile Leu Val Ser Leu Arg Gln Gln Ile Ser Glu Gly Arg
165 170 175

Arg Lys Ala Leu Ser Thr Cys Ala Ala His Leu Thr Val Val Thr Leu
180 185 190

Phe Leu Gly His Cys Ile Phe Ile Tyr Ser Arg Pro Ser Thr Ser Leu
195 200 205

Pro Glu Asp Lys Val Val Ser Val Phe Phe Thr Ala Val Thr Pro Leu
 210 215 220

Leu Asn Pro Phe Ile Tyr Ser Leu Arg Asn
 225 230 234

(6) INFORMATION POUR LA SEQ ID NO:5 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 237

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR5

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:5 :

Pro Met Tyr Leu Phe Leu Gly Asn Leu Ser Leu Leu Glu Ile Gly Tyr
 1 5 10 15

Thr Cys Ser Val Ile Pro Lys Met Leu Gln Ser Leu Val Ser Glu Ala
 20 25 30

Arg Gly Ile Ser Arg Glu Gly Cys Ala Thr Gln Met Phe Phe Phe Thr
 35 40 45

Leu Phe Ala Ile Ser Glu Cys Cys Leu Leu Ala Ala Met Ala Phe Asp
 50 55 60

Arg Tyr Met Ala Ile Cys Ser Pro Leu His Tyr Ala Thr Arg Met Ser
 65 70 75 80

Arg Gly Val Cys Ala His Leu Ala Val Val Ser Trp Thr Val Gly Cys
 85 90 95

Met Val Gly Leu Gly Gln Thr Asn Tyr Ile Phe Ser Leu Asp Phe Cys
 100 105 110

Gly Pro Cys Glu Ile Asp His Phe Phe Cys Asp Leu Pro Pro Ile Leu
 115 120 125

Ala Leu Ala Cys Gly Asp Thr Ser His Asn Glu Ala Ala Val Phe Val
 130 135 140

Val Ala Ile Leu Cys Ile Ser Ser Pro Phe Leu Leu Ile Val Ala Ser
 145 150 155 160

Tyr Gly Arg Ile Leu Ala Ala Val Leu Val Met Pro Ser Pro Glu Gly
 165 170 175

Arg Arg Lys Ala Leu Ser Thr Cys Ser Ser His Leu Leu Val Val Thr
 180 185 190

Leu Phe Tyr Gly Ser Gly Ser Val Thr Tyr Leu Arg Pro Lys Ala Ser
 195 200 205

His Ser Pro Gly Met Asp Lys Leu Leu Ala Leu Phe Tyr Thr Val Val
 210 215 220

Thr Ser Met Leu Asn Pro Ile Ile Tyr Thr Leu Arg Asn
 225 230 235 237

(7) INFORMATION POUR LA SEQ ID NO:6 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 236

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR6

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:6 :

Pro Met Tyr Leu Phe Leu Gly Asn Leu Ser Phe Leu Glu Ile Leu Tyr
 1 5 10 15

Thr Ser Thr Val Val Pro Lys Met Leu Glu Gly Phe Leu Gln Val Ala
 20 25 30

Ala Ile Ser Val Thr Gly Cys Leu Thr Gln Phe Phe Ile Phe Gly Ser
 35 40 45

Leu Ala Thr Ala Glu Cys Phe Leu Leu Ala Val Met Ala Tyr Asp Arg
 50 55 60

Phe Leu Ala Ile Cys Tyr Pro Leu Arg Tyr Pro Leu Leu Met Gly Pro
 65 70 75 80

Arg Trp Cys Met Gly Leu Val Val Thr Ala Trp Leu Ser Gly Phe Met
 85 90 95

Val Asp Glu Leu Val Val Val Leu Met Ala Gln Leu Arg Phe Cys Gly
 100 105 110

Ser Asn Arg Ile Asp His Phe Tyr Cys Asp Phe Met Pro Leu Val Val
 115 120 125

Leu Ala Cys Ser Asp Pro Arg Val Ala Gln Val Thr Thr Phe Val Leu
 130 135 140

Ser Val Val Phe Leu Thr Val Pro Phe Gly Leu Ile Leu Thr Ser Tyr
 145 150 155 160

Ala Arg Ile Val Val Thr Val Leu Arg Val Pro Ala Gly Ala Ser Arg
 165 170 175

Arg Lys Ala Phe Ser Thr Cys Ser Ser His Leu Ala Val Val Ser Thr
 180 185 190

Phe Tyr Gly Thr Leu Met Val Leu Tyr Ile Val Pro Ser Ala Val His
 195 200 205

Ser Gln Leu Leu Ser Lys Val Phe Ala Leu Leu Tyr Thr Val Val Thr
210 215 220

Pro Ile Phe Asn Pro Ile Ile Tyr Ser Phe Arg Asn
225 230 235 236

(8) INFORMATION POUR LA SEQ ID NO:7 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 237

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR7

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:7 :

Pro Arg Tyr Leu Phe Leu Gly Asn Leu Ser Leu Ala Asp Ile Gly Ile
1 5 10 15

Ser Thr Thr Thr Ile Pro Gln Met Val Val Asn Ile Gln Arg Lys Arg
20 25 30

Lys Thr Ile Ser Tyr Ala Gly Cys Leu Thr Gln Val Cys Phe Val Leu
35 40 45

Ile Phe Ala Gly Ser Glu Asn Phe Leu Leu Ala Ala Met Ala Tyr Asp
50 55 60

Arg Tyr Ala Ala Ile Cys His Pro Leu Arg Tyr Thr Ala Ile Met Asn
65 70 75 80

Pro His Leu Cys Val Leu Leu Val Met Ile Ser Leu Ser Ile Ser Thr
85 90 95

Val Asp Ala Leu Leu His Ser Leu Met Leu Leu Arg Leu Ser Phe Cys
100 105 110

Thr Asp Leu Glu Ile Pro His Phe Phe Cys Glu Leu Asp Gln Val Ile
115 120 125

Thr Leu Ala Cys Ser Asp Thr Leu Ile Asn Asn Leu Leu Ile Tyr Val
130 135 140

Thr Ala Gly Ile Phe Ala Gly Val Pro Leu Ser Gly Ile Ile Phe Ser
145 150 155 160

Tyr Leu His Ile Val Ser Ser Val Leu Arg Met Pro Ser Pro Gly Gly
165 170 175

Val Tyr Lys Ala Phe Ser Thr Cys Gly Ser His Leu Ser Val Val Cys
180 185 190

Leu Phe Tyr Gly Thr Ile Phe Gly Val Tyr Ile Ser Ser Ala Val Thr
195 200 205

Asp Ser Gln Arg Lys Gly Ala Val Ala Ser Val Met Tyr Ser Val Val
 210 215 220

Pro Gln Met Leu Asn Pro Ile Ile Tyr Thr Leu Arg Asn
 225 230 235 237

(9) INFORMATION POUR LA SEQ ID NO:8 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 176

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR8

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:8 :

Gln Ala Leu Ala Tyr Asp Arg Phe Val Ala Ile Cys His Pro Leu His
 1 5 10 15

Tyr Leu Val Ile Met Ser Pro Arg His Cys Gly Phe Leu Thr Leu Val
 20 25 30

Ser Phe Leu Leu Ser Leu Leu Asp Ser Gln Leu His Ser Phe Met Thr
 35 40 45

Leu Asn Ile Thr Ser Phe Lys Asp Val Glu Ile Ser Asn Phe Phe Cys
 50 55 60

Asp Pro Ser Gln Leu Leu Asn Leu Ser Cys Ser Asn Thr Phe Ser Asp
 65 70 75 80

Asn Ile Val Lys Tyr Phe Leu Gly Ala Phe Tyr Gly Leu Phe Pro Ile
 85 90 95

Ser Gly Ile Leu Phe Ser Tyr Tyr Lys Ile Ile Ser Ser Ile Leu Arg
 100 105 110

Ile Pro Ser Leu Gly Gly Lys Tyr Lys Ala Phe Ser Thr Cys Gly Ser
 115 120 125

His Leu Ala Val Val Cys Leu Phe Leu Val Thr Ala Ser Thr Val Tyr
 130 135 140

Leu Gly Ser Val Ala Ser His Ser Pro Arg Asn Asp Val Val Ala Ser
 145 150 155 160

Leu Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe Ile Cys Ser
 165 170 175 176

(10) INFORMATION POUR LA SEQ ID NO:9 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 176

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR9

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:9 :

Gln Ala Leu Ala Tyr Asp Arg Phe Val Ala Ile Cys Tyr Pro Leu His
 1 5 10 15
 Tyr Met Val Ile Met Asn Ser Arg Arg Cys Gly Leu Leu Ile Leu Val
 20 25 30
 Ser Trp Ile Met Ser Ala Leu His Ser Leu Leu Gln Gly Leu Met Met
 35 40 45
 Leu Arg Leu Ser Phe Cys Thr Asp Leu Glu Ile Ser His Phe Phe Cys
 50 55 60
 Glu Leu Asn His Leu Val His Leu Ala Cys Ser Asp Thr Phe Leu Asn
 65 70 75 80
 Glu Val Val Ile Tyr Phe Ala Ala Val Leu Leu Ala Gly Gly Pro Leu
 85 90 95
 Ala Gly Ile Leu Tyr Ser Tyr Cys Lys Ile Val Ser Ser Ile His Ala
 100 105 110
 Ile Ser Ser Ala Gln Gly Lys Tyr Lys Ala Phe Ser Thr Cys Ala Ser
 115 120 125
 His Leu Ser Val Val Ser Leu Phe Tyr Cys Thr Ser Pro Gly Val Tyr
 130 135 140
 Leu Ser Ser Ala Val Thr Gln Asn Ser His Ser Thr Ala Thr Ala Ser
 145 150 155 160
 Val Met Tyr Ser Val Val Thr Pro Met Leu Asn Pro Phe Ile Cys Ser
 165 170 175 176

(11) INFORMATION POUR LA SEQ ID NO:10 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 176

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR10

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:10 :

```

Gln Ala Leu Ala Tyr Asp Arg Phe Val Ala Ile Cys His Pro Leu His
1          5          10          15
Tyr Pro Arg Ile Met Ser Gln Asn Leu Cys Phe Leu Leu Val Val
20          25          30

Ser Trp Val Leu Ser Ser Ala Asn Ala Leu Leu His Thr Leu Leu Leu
35          40          45

Ala Arg Leu Ser Phe Leu Arg Gly Ile Thr Leu Pro His Phe Phe Cys
50          55          60

Asp Leu Ser Ala Leu Leu Lys Leu Ser Ser Ser Asp Thr Thr Ile Asn
65          70          75          80

Gln Leu Ala Ile Leu Thr Ala Gly Ser Ala Val Val Thr Leu Pro Phe
85          90          95

Met Cys Ile Leu Val Ser Tyr Gly His Ile Gly Ala Thr Ile Leu Arg
100         105         110

Arg Pro Ser Leu Lys Gly Ile Cys Lys Ala Leu Ser Thr Cys Gly Ser
115         120         125

His Leu Ser Val Val Ser Val Tyr Tyr Gly Ala Val Ile Ala Leu Tyr
130         135         140

Ile Val Pro Ser Ser Asn Ser Thr Asn Asp Lys Asp Ile Ala Val Ser
145         150         155         160

Val Leu Tyr Thr Leu Val Ile Pro Met Leu Asn Pro Phe Ile Cys Ser
165         170         175 176

```

(12) INFORMATION POUR LA SEQ ID NO:11 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 176

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR11

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:11 :

```

Gln Ala Leu Ala Tyr Asp Arg Phe Leu Ala Ile Cys Tyr Pro Leu His
1          5          10          15
Tyr Thr Val Ile Met Asn Pro Arg Leu Cys Gly Phe Ser Ile Leu Val
20          25          30

Ser Phe Leu Leu Ser Leu Leu Asp Ser Gln Leu His Asn Leu Met Ile
35          40          45

Leu Gln Ile Thr Ser Phe Lys Asp Val Glu Ile Ser Ser Phe Phe Cys
50          55          60

Asp Pro Ser Gln Leu Leu Asn Leu Ser Cys Ser Asp Asn Tyr Ser Ile
65          70          75          80

```

Asn Thr Gly Lys Tyr Val Leu Phe Ala Leu Tyr Ser Phe Phe Pro Ile
 85 90 95

Ser Gly Ile Leu Phe Ser Tyr Tyr Lys Ile Ile Ser Ser Ile Leu Arg
 100 105 110

Ile Pro Ser Ser Gly Gly Lys Tyr Lys Ala Phe Ser Thr Cys Gly Ser
 115 120 125

His Leu Ala Val Phe Cys Leu Phe Leu Gly Thr Gly Thr Ala Val Tyr
 130 135 140

Phe Gly Ser Ala Val Ser His Ser Pro Arg Glu Asn Val Val Ser Ser
 145 150 155 160

Val Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe Ile Cys Ser
 165 170 175 176

(13) INFORMATION POUR LA SEQ ID NO:12 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 176

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR12

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:12 :

Gln Ala Leu Ala Tyr Asp Arg Phe Val Ala Ile Cys His Pro Leu His
 1 5 10 15

Tyr Pro Thr Ile Met Asn Pro Arg Phe Cys Gly Phe Leu Val Leu Val
 20 25 30

Ser Phe Leu Val Ser Leu Leu Glu Ser Gln Leu His Asn Leu Ile Ala
 35 40 45

Leu Gln Phe Thr Thr Phe Lys Asp Val Lys Ile Ala Asn Phe Phe Cys
 50 55 60

Asp Pro Ser Gln Val Leu Ser Leu Ser Cys Ser Gly Thr Phe Ile Asn
 65 70 75 80

Ile Ile Val Met Tyr Phe Val Gly Ala Leu Phe Gly Val Phe Pro Ile
 85 90 95

Ser Gly Ile Leu Phe Ser Tyr Tyr Lys Ile Val Ser Thr Ile Leu Arg
 100 105 110

Ile Pro Ser Ser Gly Gly Lys Tyr Lys Ala Phe Ser Thr Cys Gly Ser
 115 120 125

His Leu Ser Val Val Cys Leu Phe Tyr Gly Thr Gly Phe Gly Val Tyr
 130 135 140

Leu Gly Ser Ala Val Ser His Ser Ser Arg Lys Ser Ala Val Ala Ser
145 150 155 160

Val Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe Ile Cys Ser
165 170 175 176

(14) INFORMATION POUR LA SEQ ID NO:13 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 168

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR13

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:13 :

Gly Ala Ile Arg Phe Pro Leu His Asn Thr Thr Ile Met Ser Pro Lys
1 5 10 15

Leu Gly Leu Phe Leu Val Val Leu Ser Trp Val Leu Thr Met Phe His
20 25 30

Ala Met Leu His Thr Leu Leu Met Ala Arg Leu Cys Phe Cys Ala Glu
35 40 45

Asn Met Ile Pro His Phe Phe Cys Asp Met Ser Ala Leu Leu Lys Leu
50 55 60

Ser Cys Ser Asn Thr His Val Asn Glu Leu Val Ile Phe Ile Thr Ala
65 70 75 80

Gly Leu Ile Leu Leu Ile Pro Leu Val Leu Ile Leu Leu Ser Tyr Gly
85 90 95

His Ile Val Ser Ser Ile Leu Lys Val Pro Ser Ala Arg Gly Ile His
100 105 110

Lys Thr Phe Ser Thr Cys Gly Ser His Leu Ser Val Val Ser Leu Phe
115 120 125

Tyr Gly Thr Val Ile Gly Leu Tyr Leu Cys Pro Ser Ala Asn Asn Ser
130 135 140

Thr Val Lys Asp Thr Val Met Ala Leu Met Tyr Thr Val Val Thr Pro
145 150 155 160

Met Leu Asn Pro Phe Ile Cys Ser
165 168

(15) INFORMATION POUR LA SEQ ID NO:14 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 176

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR14

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:14 :

Gln Ala Leu Ala Tyr Asp Arg Phe Leu Ala Ile Cys His Pro Leu His
 1 5 10 15
 Tyr Thr Ala Ile Met Asn Pro Arg Leu Cys Gly Leu Leu Val Leu Val
 20 25 30
 Cys Trp Ile Leu Ser Val Leu His Ala Leu Leu Gln Ser Leu Met Val
 35 40 45
 Leu Arg Leu Ser Phe Cys Arg Asp Ile Glu Ile Pro His Phe Phe Cys
 50 55 60
 Glu Leu Asn Gln Val Val Gln Leu Ala Cys Phe Asp Asn Leu Leu Asn
 65 70 75 80
 Asp Ile Val Met Asn Phe Ala Leu Val Leu Leu Ala Thr Cys Pro Leu
 85 90 95
 Ala Gly Ile Leu Tyr Ser Tyr Ser Lys Ile Val Ser Ser Ile Arg Ala
 100 105 110
 Ile Ser Ser Ala Gln Gly Lys Tyr Lys Ala Phe Ser Thr Cys Ala Ser
 115 120 125
 His Leu Ser Val Val Ser Leu Phe Tyr Cys Thr Gly Leu Gly Val Tyr
 130 135 140
 Leu Ser Ser Ala Val Ser His Ser Ser Arg Ser Ser Ala Thr Ala Ser
 145 150 155 160
 Val Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe Ile Cys Ser
 165 170 175 176

(16) INFORMATION POUR LA SEQ ID NO:15 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 119

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR15

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:15 :

His Leu Cys Arg Leu His Leu Thr Val Leu Lys Leu Ala Cys Ser Asp
 1 5 10 15
 Thr Leu Ile Asn Asn Ile Val Val Phe Ser Met Ile Ile Val Leu Gly
 20 25 30
 Val Phe Pro Leu Ser Gly Ile Leu Phe Ser Tyr Ser Gln Ile Phe Ser
 35 40 45
 Ser Ile Leu Arg Ile Ser Ser Asp Arg Gly Lys Tyr Lys Val Phe Ser
 50 55 60
 Thr Cys Gly Ser His Leu Leu Val Val Ser Leu Phe Tyr Gly Ser Ser
 65 70 75 80
 Leu Gly Val Tyr Leu Ser Ser Val Ala Thr Leu Ser Ser Arg Met Thr
 85 90 95
 Leu Met Ala Ser Val Met Tyr Thr Met Val Thr Pro Met Leu Asn Pro
 100 105 110
 Ile Ile Tyr Thr Leu Arg Asn
 115 119

(17) INFORMATION POUR LA SEQ ID NO:16 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 159

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR16

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:16 :

Pro Met Tyr Leu Phe Leu Gly Asn Leu Ser Phe Leu Glu Ile Leu Tyr
 1 5 10 15
 Thr Ser Thr Val Val Pro Lys Met Leu Glu Gly Phe Leu Gln Val Ala
 20 25 30
 Ala Ile Ser Val Thr Gly Cys Leu Thr Gln Phe Phe Ile Phe Gly Ser
 35 40 45
 Leu Ala Thr Ala Glu Cys Phe Leu Leu Ala Val Met Ala Tyr Asp Arg
 50 55 60
 Phe Leu Ala Ile Cys Tyr Pro Leu Arg Tyr Pro Leu Leu Met Gly Pro
 65 70 75 80
 Arg Trp Cys Met Gly Leu Val Val Thr Ala Trp Leu Ser Gly Phe Met
 85 90 95
 Val Asp Glu Leu Val Val Val Leu Met Ala Gln Leu Arg Phe Cys Gly
 100 105 110
 Ser Asn Arg Ile Asp His Phe Tyr Cys His Phe Met Pro Leu Val Val
 115 120 125

Leu Ala Cys Ser Asp Pro Arg Val Ala Gln Val Thr Thr Phe Val Leu
 130 135 140

Ser Val Val Pro Leu Thr Val Pro Phe Gly Leu Ile Leu Thr Ser
 145 150 155 159

(18) INFORMATION POUR LA SEQ ID NO:17 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 113
- (B) TYPE: acides aminés
- (ii) TYPE DE MOLECULE: protéine
- (ix) CARACTERISTIQUES
- (A) NOM/CLE: AMOR17

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:17 :

Glu Asp Leu Cys Ala Arg Leu Lys Arg Ser Arg Ser Asp Thr Thr Ile
 1 5 10 15

Asn Glu Val Gly Ile Leu Thr Ala Gly Ser Ala Val Val Thr Leu Pro
 20 25 30

Phe Met Cys Ile Leu Val Ser Tyr Gly His Met Gly Ala Thr Ile Leu
 35 40 45

Arg Arg Pro Ser Leu Lys Gly Met Cys Lys Ala Leu Ser Thr Cys Gly
 50 55 60

Ser His Leu Cys Val Val Ser Val Tyr Tyr Gly Ala Val Ile Ala Leu
 65 70 75 80

Tyr Ile Val Pro Ser Ser Asn Ser Thr Asn Asp Lys Asp Ile Ala Val
 85 90 95

Ser Val Leu Tyr Thr Leu Val Ile Pro Met Leu Asn Pro Phe Ile Cys
 100 105 110

Ser
 113

(19) INFORMATION POUR LA SEQ ID NO:18 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 176
- (B) TYPE: acides aminés
- (ii) TYPE DE MOLECULE: protéine
- (ix) CARACTERISTIQUES
- (A) NOM/CLE: AMOR18

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:18 :

```

Gln Ala Leu Gly Tyr Asp Arg Phe Val Ala Met Cys His Pro Gly Gln
1          5          10          15

Tyr Leu Val Ile Met Ser Pro Arg His Gly Gly Phe Leu Thr Leu Val
20          25          30

Ser Phe Leu Leu Ser Leu Leu Asp Ser Gln Leu His Ser Phe Met Thr
35          40          45

Leu Asn Ile Thr Ser Phe Lys Asp Val Glu Ile Ser Asn Phe Phe Cys
50          55          60

Asp Pro Ser Gln Leu Leu Asn Leu Ser Cys Ser Asn Thr Phe Ser Asp
65          70          75          80

Asn Ile Val Lys Tyr Phe Leu Gly Ala Phe Tyr Gly Leu Phe Pro Ile
85          90          95

Ser Gly Ile Leu Phe Ser Tyr Tyr Lys Ile Ile Ser Ser Ile Leu Arg
100         105         110

Ile Pro Ser Leu Gly Gly Lys Tyr Lys Ala Phe Ser Thr Cys Gly Ser
115         120         125

His Leu Ala Val Val Cys Leu Phe Leu Val Thr Ala Ser Thr Val Tyr
130         135         140

Leu Gly Ser Val Ala Ser His Ser Pro Arg Asn Asp Val Val Ala Ser
145         150         155         160

Leu Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe Ile Cys Ser
165         170         175 176

```

(20) INFORMATION POUR LA SEQ ID NO:19 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 176

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR19

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:19 :

```

Gln Ala Leu Ala Tyr Asp Arg Phe Leu Ala Ile Cys His Pro Leu His
1          5          10          15

Tyr Leu Val Ile Met Ser Pro Arg His Cys Gly Phe Leu Thr Leu Val
20          25          30

Ser Phe Leu Leu Ser Leu Leu Asp Ser Gln Leu His Ser Phe Met Thr
35          40          45

Leu Asn Ile Thr Ser Phe Lys Asp Val Glu Ile Ser Asn Phe Phe Cys
50          55          60

Asp Pro Ser Gln Leu Leu Asn Leu Ser Cys Ser Asn Thr Phe Ser Asp
65          70          75          80

```

Asn Ile Val Lys Tyr Phe Leu Gly Ala Phe Tyr Gly Leu Phe Pro Ile
85 90 95

Ser Gly Ile Leu Phe Ser Tyr Tyr Lys Ile Ile Ser Ser Ile Leu Arg
100 105 110

Ile Pro Ser Leu Gly Gly Lys Tyr Lys Ala Phe Ser Thr Cys Gly Ser
115 120 125

His Leu Ala Val Val Cys Leu Phe Leu Val Thr Ala Ser Thr Val Tyr
130 135 140

Leu Gly Ser Val Ala Ser His Ser Pro Arg Asn Asp Val Val Ala Ser
145 150 155 160

Leu Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe Ile Cys Ser
165 170 175 176

(21) INFORMATION POUR LA SEQ ID NO:20 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 176

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR20

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:20 :

Gln Ala Leu Ala Tyr Asp Arg Phe Leu Ala Ile Cys His Pro Arg His
1 5 10 15

Tyr Leu Val Ile Met Ser Pro Arg His Cys Gly Phe Leu Thr Leu Val
20 25 30

Ser Phe Leu Leu Ser Leu Leu Asp Ser Gln Leu His Ser Phe Met Thr
35 40 45

Leu Asn Ile Thr Ser Phe Lys Asp Val Glu Ile Ser Asn Phe Phe Cys
50 55 60

Asp Pro Ser Gln Leu Leu Asn Leu Ser Cys Ser Asn Thr Phe Ser Asp
65 70 75 80

Asn Ile Val Lys Tyr Phe Leu Gly Ala Phe Tyr Gly Leu Phe Pro Ile
85 90 95

Ser Gly Ile Leu Phe Ser Tyr Tyr Lys Ile Ile Ser Ser Ile Leu Arg
100 105 110

Ile Pro Ser Leu Gly Gly Lys Tyr Lys Ala Phe Ser Thr Cys Gly Ser
115 120 125

His Leu Ala Val Val Cys Leu Phe Leu Val Thr Ala Ser Thr Val Tyr
130 135 140

Leu Gly Ser Val Ala Ser His Ser Pro Arg Asn Asp Val Val Ala Ser
145 150 155 160

Leu Met Tyr Thr Val Val Thr Pro Met Leu Asn Pro Phe Ile Cys Ser
165 170 175 176

(22) INFORMATION POUR LA SEQ ID NO:21 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 112

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR21

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:21 :

Pro Met Tyr Leu Phe Leu Gly Asn Leu Ser Phe Met Asp Ile Cys Phe
1 5 10 15

Thr Thr Val Val Val Pro Lys Met Leu Ala Asn Leu Leu Ser Glu Thr
20 25 30

Lys Gly Ile Ser Tyr Val Gly Cys Leu Val Gln Met Tyr Phe Phe Met
35 40 45

Ala Phe Gly Asn Thr Asp Ser Tyr Leu Leu Ala Ser Met Ala Ile Asp
50 55 60

Arg Leu Val Ala Ile Cys Asn Pro Leu His Tyr Asp Val Ala Met Arg
65 70 75 80

Pro His Arg Cys Leu Leu Met Leu Leu Gly Ser Cys Thr Ile Ser His
85 90 95

Leu His Ala Leu Phe Arg Val Leu Leu Met Ser Arg Leu Ser Phe Cys
100 105 110 112

(23) INFORMATION POUR LA SEQ ID NO:22 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 119

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR22

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:22 :

His Leu Cys Arg Leu His Leu Thr Val Leu Lys Leu Ala Cys Ser Asp
1 5 10 15

Thr Leu Ile Asn Asn Ile Val Val Phe Ser Met Ile Ile Val Leu Gly
 20 25 30
 Val Phe Pro Leu Ser Gly Ile Leu Phe Ser Tyr Ser Gln Ile Phe Ser
 35 40 45
 Ser Ile Leu Arg Ile Ser Ser Asp Arg Gly Lys Tyr Lys Val Phe Ser
 50 55 60
 Thr Cys Gly Ser His Leu Leu Val Val Ser Leu Phe Tyr Gly Ser Ser
 65 70 75 80
 Leu Gly Val Tyr Leu Ser Ser Val Ala Thr Leu Ser Ser Arg Met Thr
 85 90 95
 Leu Met Ala Ser Val Met Tyr Thr Met Val Thr Pro Met Leu Asn Pro
 100 105 110
 Ile Ile Tyr Thr Leu Arg Asn
 115 119

(24) INFORMATION POUR LA SEQ ID NO:23 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 141

(B) TYPE: acides aminés

(ii) TYPE DE MOLECULE: protéine

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR23

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:23 :

Trp Ser Leu Leu Glu Ser Gln Leu His Ser Leu Arg Thr Leu Asn Met
 1 5 10 15
 Thr Ser Phe Arg Asp Val Glu Ser Ser Asn Leu Leu Cys Asp Pro Ser
 20 25 30
 Gln Met Leu Asn Leu Ser Cys Ser Asn Thr Phe Ser Asp Asn Ile Val
 35 40 45
 Lys Tyr Phe Leu Gly Ala Phe Tyr Gly Leu Phe Pro Ile Ser Gly Ile
 50 55 60
 Leu Phe Ser Tyr Tyr Lys Ile Ile Ser Ser Ile Leu Arg Ile Pro Ser
 65 70 75 80
 Leu Gly Gly Lys Tyr Lys Ala Phe Ser Thr Cys Gly Ser His Leu Ala
 85 90 95
 Val Val Cys Leu Phe Leu Val Thr Ala Ser Thr Val Tyr Leu Gly Ser
 100 105 110
 Val Ala Ser His Ser Pro Arg Asn Asp Val Val Ala Ser Leu Met Tyr
 115 120 125

Thr Val Val Thr Pro Met Leu Asn Pro Phe Ile Cys Ser
130 135 140 141

(25) INFORMATION POUR LA SEQ ID NO:24 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 711
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 1

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:24 :

CCT ATG TAT TTG TTC CTT GGC AAC TTG TCC TTC CTG GAC CTC AGC TTC	48
ACC ACC AGC TCC ATC CCC CAG CTG CTC CAC AAC CTG AGT GGC CGT GAC	96
AAG ACC ATC AGC TAT GTG GGC TGC GTG GTC CAG CTC TTC CTG TTC CTG	144
GGC CTG GGT GGA GTG GAG TGT CTA CTG CTG GCC GTC ATG GCC TAT GAC	192
AGG TTC GTG GCC GTC TGC AAG CCC CTG CAC TAC ACG GTG ATC ATG AGT	240
TCC AGG CTC TGC CTG GGC TTG GTG TCA GTG GCC TGG GGC TGT GGA ATG	288
GCC AAC TCC TTG GTC ATG TCT CCA GTG ACC CTA CAA TTA CCC CGC TGC	336
GGG CAC AAT AAG GTG GAC CAT TTC CTG TGT GAG ATG CCA GCC CTG ATC	384
CGC ATG GCC TGC GTC AAC ACA GTG GCC ATA GAA GGC ACT GTC TTT GTC	432
CTG GCC GTG GGC ATC GTG CTG TCT CCC CTG GTC TTC ATC TTG GTG TCC	480
TAT GGC CAC ATC GTC AGG GCG GTG TTC AGA ATC CAG TCG TCC TCA GGA	528
AGA CAC AGA ATC TTC AAC ACC TGT GGC TCC CAC CTC ACC GTG GTC TCC	576
CTG TTC TAC GGG AAC ATC ATC TAC ATG TAC ATG CAG CCA GGA AGC AGG	624
TCC TCC CAG GAC CAG GGC AAG TTC CTC ACC CTC TTC TAC AAC ATC GTC	672
ACC CCC CTC CTG AAC CCC TTC ATC TAT TCC CTC AGG AAT	711

(26) INFORMATION POUR LA SEQ ID NO:25 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 711
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 2

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:25 :

CCC ATG TAT TTG TTC CTT GGT AAC CTG TCC TTT GTG GAA GTC TGT TTA	48
ACC TCC ACC ACG GTC CCC AAG ATA CTG GTG AAC ACG CAG ACA CTC AGC	96
AAA GAC ATC TCC TAC AGA GGC TGC CTT ACT CAG GTG TAT TTT TTA ATG	144
GTT TTT GCA GGT ATG GAT AAT TTC CTT CTG ACT GTG ATG GCC TTT GAC	192
CGC TTT GTG GCC ATC TGC TAC CCC CTG AAC TAT ACG GTC ATC ATG AAC	240
CCC AGG CTC TGT GTC CTC CTG GTG CTG CTG TCT TGG CTC ATC ATG TTC	288
TGG GTG TCC TTA CTT CAC ATT CTA CTC CTG AAG CGA CTG ACC TTC TCC	336
AGT GGC ACT GCA GTC CCT CAT TTC TGT GAA CTG TCT CAG CTT CTC	384
AAA GCA ACC AGC TCT GAC ACC CTC GTC AAT ATC ATC TTA CTG TAT GTG	432
GTG ACT GCC CTG CTG GGT ATC TTC CCT GCC ACT GGG ATC CTC TAC TCC	480
TAC TCT CAG ATC GTC TCT TCC TTA CTG AGG ATG TCC TCC TCT GTG GGC	528
AAG TCT AAA GCC TTC TCC ACC TGT GGT TCC CAC CTC TGT GTG GTC TCC	576

TTG TTC TAT GGA ACA GGT CTT GGG GTT CAC CTC AGT TCT GCC ATG AAC	624	-
CAT CCT TCT CAG GGA AAC ATG ATT GCC TCC GTG ATG TTA CAC TGT GGT	672	
CAC CCC ATG CTG AAC CCC ATC ATC TAC ACC CTC CGG AAC	711	

(27) INFORMATION POUR LA SEQ ID NO:26 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 711
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 3

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:26 :

CCC ATG TAC TTG TTT CTT GGC AAT CTG TCC TTC CTG GAG GTC TGG TAC	48
ACC ACG GCC GCA GTG CCC AAA GCC CTG GCC ATC CTG CTG GGG AGG AGC	96
CAG AGC ATC TCC TTC ATC AGC TGC CTC CTG CAG ATG TAC CTG GTC TTC	144
TCG CTG GGC TGC ACG GAG TAC TTC CTC CTT GTG GCC ATG GCT TAT GAC	192
CGC TAT GTG GCC ATC TGC TTC CCC CTG CAC TAC ACC ACC ATC ATG AGC	240
CTC AAG CTC TGT CTC TCC CTG GTG GTG CTG TCC TGG GTG CTG ACC ATG	288
CTC CAT GCC CTG TTG CAC ACT CTG CTT GTG GTC AGA TTG TCT TTC TGT	336
TCG GAC AAT GTA ATC CCA CAC TTT TCC TGT GAA ATA TCT GCT TTA TTG	384
AAG CTG GCC TGC TCC AAC ACT CAT GTC AAT GAA CTG GTG ATA TTT ATC	432
ACG GGA GGA CTT GTT ATT GTC ACC CCA TTT CTA CTC ATC CTT GGG TCC	480
TAT GTA CAA ATT TTC TCC TCC ATC CTC AAG GTC CCT TCT GCT CGT GGT	528
ATC CAC AAG GCC TTC TCT ACC TGT GGC TCC CAC CTC TCT GTG GTG TCA	576
CTG TTC TAT GGG ACA ATT ATT GGT CTC TAT TTA TGT CCA TCA GCT AAT	624
AAT TCT ACT GTG AAA GAC ACT GTC GTG GCT CTG ATG TAC ACG GTG GTG	672
ACT CCC ATG CTG AAC CCC TTC ATC TAC ACC CTC CGA AAT	711

(28) INFORMATION POUR LA SEQ ID NO:27 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 702
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 4

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:27 :

CCC ATG TAC TTG TTT CTC GGT AAC CTG TCC TTT ATC GAT GTC TGC CAC	48
TCC ACT GTC ACT GTG CCC AAG ATG CTG AGA GAC ACC TGG TCA GAG GAG	96
AAG CTC ATC TCC TTT GAT GCC TGT GTG ACC CAG ATG TTC TTC CTG CAC	144
CTC TTT GCC TGC ACA GAG ATC TTC CTC CTC ACC GTC ATG GCC TAT GAT	192
CGT TAT GTG GCC ATC TGT AAA CCC CTG CAG TAC ATG ACA GTG ATG AAT	240
TGG AAG GTA TGT GTG CTG CTG GCT GTG GCC CTC TGG GCA GGA GGA ACC	288

ATC CAC TCC ATA TCC CTG ACC TCC CTC ACC ATC AAG CTG CCC TAC TGT	336
GGT CCT GAT GAG ATT GAC AAC TTC TTC TGT GAC GTG CCG CAG GTG ATC	384
AAA TTG GCC TGC ACT GAC ACC CAC ATC ATT GAG ATC CTC ATC GTC TCC	432
AAC AGT GGG CTG ATC TCC GTG GTC TGT TTT GTC GTC CTT GTG GTG TCC	480
TAT GCA GTC ATC CTG GTG AGT CTG CGG CAG CAG ATC TCC GAG GGC AGG	528
CGG AAG GCC CTG TCC ACC TGT GCA GCC CAC CTC ACT GTG GTC ACA CTG	576
TTC CTG GGA CAC TGC ATC TTC ATC TAT TCC CGC CCA TCC ACC AGC CTC	624
CCA GAG GAC AAA GTG GTG TCT GTG TTT TTC ACT GCT GTC ACC CCT CTG	672
CTA AAC CCC TTC ATC TAC TCC CTC CGA AAT	702

(29) INFORMATION POUR LA SEQ ID NO:28 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 711
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 5

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:28 :

CCC ATG TAT TTG TTC CTT GGT AAC TTG TCT CTC CTA GAG ATT GGC TAC	48
ACT TGC TCT GTC ATA CCC AAG ATG CTG CAG AGT CTT GTG AGT GAG GCC	96
CGA GGA ATC TCT CGG GAG GGT TGT GCC ACA CAG ATG TTT TTC TTT ACA	144
TTA TTT GCT ATC AGT GAG TGC TGC CTT TTG GCA GCC ATG GCT TTT GAC	192
CGC TAT ATG GCC ATA TGT TCC CCA CTC CAC TAT GCA ACA CGA ATG AGT	240
CGT GGG GTG TGT GCC CAT TTG GCA GTG GTT TCT TGG ACA GTG GGA TGC	288
ATG GTA GGC TTG GGC CAA ACC AAT TAT ATT TTC TCC TTA GAC TTC TGT	336
GGC CCC TGT GAG ATA GAC CAC TTC TTC TGT GAT CTC CCA CCT ATC CTG	384
GCA CTT GCT TGT GGG GAT ACA TCC CAT AAT GAG GCT GCG GTC TTT GTG	432
GTG GCA ATC CTT TGC ATT TCT AGC CCA TTT TTA TTG ATC GTT GCT TCC	480
TAT GGC AGA ATT CTA GCT GCA GTC CTG GTC ATG CCC TCC CCT GAG GGC	528
CGG CGG AAA GCT CTC TCC ACC TGT TCT TCC CAC TTA CTT GTA GTA ACG	576
CTC TTC TAT GGC TCA GGA TCT GTT ACC TAC TTG AGG CCC AAG GCT AGC	624
CAC TCA CCA GGA ATG GAT AAA CTG CTA GCC CTC TTC TAT ACC GTG GTG	672
ACA TCC ATG CTC AAC CCC ATC ATC TAC ACC CTC CGG AAC	711

(30) INFORMATION POUR LA SEQ ID NO:29 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 708
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 6

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:29 :

CCC ATG TAC TTG TTC CTC GGT AAT TTG TCC TTC CTG GAG ATC CTT TAT	48
---	----

ACA TCC ACA GTG GTG CCG AAA ATG CTG GAG GGC TTC CTG CAG GTG GCA	96
GCC ATC TCT GTG ACT GGT TGC TTG ACC CAG TTC TTC ATC TTT GGT TCT	144
CTA GCC ACA GCA GAA TGC TTC CTA CTG GCT GTT ATG GCA TAT GAT CGC	192
TTC TTG GCA ATC TGC TAC CCA CTT CGC TAT CCA CTC CTG ATG GGG CCT	240
AGA TGG TGC ATG GGG CTG GTG GTC ACA GCC TGG CTG TCT GGC TTC ATG	288
GTA GAT GAA TTA GTT GTG GTC CTG ATG GCC CAG CTG AGG TTC TGT GGC	336
TCC AAT CGC ATC GAT CAC TTT TAC TGT GAC TTC ATG CCT TTG GTG GTC	384
CTG GCT TGC TCA GAT CCC AGA GTA GCC CAG GTG ACA ACA TTT GTT CTC	432
TCT GTA GTC TTC CTC ACT GTT CCA TTT GGA CTG ATT CTG ACA TCC TAT	480
GCT CGC ATC GTG GTG ACT GTG CTG AGA GTT CCT GCT GGG GCC AGC AGG	528
AGA AAG GCT TTT TCC ACA TGC TCC CAC CTT GCT GTA GTG TCC ACC	576
TTC TAT GGA ACT CTC ATG GTC TTG TAC ATT GTG CCC TCA GCT GTC CAC	624
TCC CAG CTC CTC TCC AAG GTC TTT GCC TTG CTC TAT ACT GTG GTC ACT	672
CCC ATC TTC AAC CCC ATC ATC TAC TCC TTC CGG AAT	708

(31) INFORMATION POUR LA SEQ ID NO:30 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 711
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 7

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:30 :

CCC AGG TAC TTG TTT CTT GGC AAT TTG TCT TTG GCC GAC ATT GGG ATC	48
AGC ACA ACC ACG ATC CCC CAG ATG GTG AAC ATC CAG AGA AAG AGA	96
AAG ACC ATC AGT TAC GCA GGC TGC CTC ACC CAG GTC TGC TTT GTC CTG	144
ATT TTT GCT GGA TCG GAG AAC TTT CTC CTT GCA GCA ATG GCT TAT GAC	192
CGT TAC GCA GCC ATC TGC CAT CCC CTC AGG TAC ACG GCC ATC ATG AAC	240
CCC CAC CTG TGT GTC CTG CTG GTT ATG ATC TCC TTG TCC ATC AGC ACG	288
GTG GAT GCC CTG CTG CAC AGT CTG ATG CTG AGG CTG TCC TTC TGC	336
ACA GAC CTG GAG ATC CCC CAC TTC TTC TGT GAA CTT GAT CAG GTG ATC	384
ACA CTG GCC TGT TCT GAC ACC CTC ATC AAT AAC CTC CTG ATA TAT GTC	432
ACA GCT GGG ATA TTT GCT GGT GTT CCT CTC TCT GGA ATC ATC TTC TCT	480
TAC CTT CAC ATT GTG TCC TCT GTC TTG AGA ATG CCA TCA CCA GGA GGA	528
GTG TAT AAA GCC TTT TCC ACC TGT GGC TCT CAC CTG TCT GTG GTC TGC	576
TTG TTC TAT GGG ACA ATT TTT GGG GTG TAC ATT AGC TCT GCA GTG ACT	624
GAC TCA CAG AGA AAA GGT GCA GTG GCC TCA GTG ATG TAC TCT GTG GTC	672
CCT CAG ATG CTG AAC CCC ATC ATC TAC ACC CTC AGA AAC	711

(32) INFORMATION POUR LA SEQ ID NO:31 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 528
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR 8

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:31 :

CAA GCT TTG GCG TAT GAC AGG TTT GTG GCC ATC TGT CAT CCT CTG CAT	48
TAT CTG GTC ATT ATG AGC CCT CGC CAT TGT GGC TTC TTA ACT TTG GTG	96
TCA TTT TTG CTG AGT CTT TTG GAC TCC CAG CTG CAC AGT TTC ATG ACC	144
TTA AAT ATT ACC AGC TTC AAG GAT GTG GAA ATT TCT AAT TTC TTC TGT	192
GAC CCT TCT CAA CTG CTG AAT CTC TCC TGT TCC AAC ACC TTC TCT GAT	240
AAC ATT GTC AAG TAT TTT CTG GGA GCC TTC TAT GGC CTT TTT CCC ATC	288
TCA GGG ATC CTT TTC TCT TAC TAC AAA ATT ATT TCC TCC ATT CTG AGG	336
ATC CCC TCC TTA GGT GGG AAG TAC AAA GCC TTC TCC ACC TGT GGG TCT	384
CAC CTG GCA GTT GTT TGC TTA TTT TTA GTG ACA GCC TCC ACA GTG TAC	432
CTT GGA TCA GTT GCA TCA CAT TCT CCC AGA AAT GAT GTG GTG GCT TCT	480
CTG ATG TAC ACT GTG GTC ACC CCC ATG CTC AAT CCC TTC ATC TGC AGT	528

(33) INFORMATION POUR LA SEQ ID NO:32 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 528

(B) TYPE: acide nucléique

(C) NOMBRE DE BRIN: double

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR 9

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:32 :

CAA GCT TTG GCG TAT GAT AGG TTT GTG GCC ATC TGC TAC CCC CTG CAC	48
TAC ATG GTC ATC ATG AAC TCC CGG CGA TGT GGA TTG CTG ATT CTG GTG	96
TCT TGG ATC ATG AGT GCT CTT CAT TCC TTG TTA CAA GGT TTA ATG ATG	144
TTG AGA CTG TCC TTC TGC ACA GAT TTG GAA ATC TCC CAC TTT TTC TGT	192
GAA CTT AAT CAC CTG GTC CAT CTT GCC TGC TCT GAC ACC TTT CTC AAT	240
GAG GTG GTG ATA TAT TTT GCT GCT GTC TTG CTG GCT GGT GGC CCC CTC	288
GCT GGC ATC CTT TAC TCT TAC TGC AAG ATA GTC TCC TCC ATC CAT GCA	336
ATC TCT TCA GCT CAG GGC AAG TAC AAA GCC TTC TCC ACC TGT GCA TCT	384
CAC CTC TCC GTG GTC TCC TTA TTT TAT TGT ACA AGC CCG GGT GTG TAC	432
CTC AGT TCT GCT GTG ACC CAA AAC TCA CAC TCC ACT GCA ACT GCC TCG	480
GTG ATG TAC AGC GTG GTC ACC CCC ATG CTC AAC CCC TTT ATC TGC AGT	528

(34) INFORMATION POUR LA SEQ ID NO:33 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 528

(B) TYPE: acide nucléique

(C) NOMBRE DE BRIN: double

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

(A) NOM/CLE: AMOR 10

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:33 :

CAA GCT TTG GCG TAC GAC AGG TTT GTG GCC ATC TGT CAC CCA CTG CAT	48
TAT CCC AGA ATC ATG AGT CAG AAC CTC TGT TTC CTG CTA GTG GTT GTG	96
TCC TGG GTC TTA TCC TCT GCC AAT GCC CTT TTG CAC ACC CTC CTC CTA	144
GCC CGT CTC TCT TTC CTT AGA GGC ATC ACT CTG CCC CAC TTC TTC TGT	192
GAT CTC TCT GCG TTA CTC AAG CTA TCC AGC TCT GAC ACC ACC ATC AAT	240
CAG CTG GCT ATT CTC ACG GCA GGA TCA GCA GTT GTT ACC CTG CCA TTC	288
ATG TGC ATT CTG GTC TCA TAT GGC CAC ATT GGG GCC ACC ATC CTG AGA	336
AGA CCC TCC CTC AAG GGC ATC TGC AAA GCC TTA TCC ACA TGT GGC TCC	384
CAC CTC TCT GTG TAC TAT GGA GCA GTT ATT GCA CTC TAT	432
ATT GTC CCC TCA TCT AAT AGC ACT AAT GAC AAG GAT ATT GCT GTG TCT	480
GTG TTG TAT ACT CTG GTC ATC CCC ATG CTC AAC CCC TTC ATC TGC AGT	528

(35) INFORMATION POUR LA SEQ ID NO:34 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 528
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 11

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:34 :

CAA GCT TTG GCG TAT GAT AGG TTC TTG GCC ATC TGT TAT CCC CTG CAT	48
TAT ACA GTC ATT ATG AAC CCT CGC CTC TGT GGC TTC TCA ATT TTG GTA	96
TCA TTT TTG CTC AGT TTC TTG GAC TCT CAG CTG CAC AAT TTG ATG ATC	144
TTA CAA ATT ACC AGT TTC AAG GAT GTG GAA ATT TCT AGT TTC TTC TGT	192
GAC CCT TCT CAA CTT CTG AAT CTT TCC TGT TCT GAC AAC TAC TCT ATT	240
AAT ACT GGC AAG TAT GTT CTT TTT GCC CTA TAT AGC TTT TTC CCC ATC	288
TCA GGG ATC CTT TTC TCT TAC TAT AAA ATA ATT TCC TCC ATT CTG AGG	336
ATC CCA TCC TCA GGG GGG AAG TAC AAA GCC TTC TCC ACT TGT GGC TCT	384
CAC CTG GCA GTT TTT TGC CTA TTT TTA GGA ACA GGT ACT GCA GTG TAC	432
TTT GGA TCA GCT GTA TCA CAT TCT CCC AGG GAG AAT GTG GTG TCC TCA	480
GTA ATG TAT ACT GTG GTC ACC CCC ATG CTC AAT CCC TTT ATC TGC AGT	528

(36) INFORMATION POUR LA SEQ ID NO:35 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 528
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 12

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:35 :

CAA GCT TTG GCG TAT GAC AGG TTT GTG GCC ATC TGT CAC CCC CTG CAT	48
TAT CCA ACC ATT ATG AAC CCT CGA TTT TGT GGC TTT TTA GTT TTG GTG	96
TCT TTT TTG GTT AGC CTT TTG GAA TCC CAG CTG CAC AAT TTG ATT GCA	144
TTA CAG TTT ACT ACT TTC AAA GAT GTA AAA ATT GCT AAT TTT TTC TGT	192
GAC CCT TCT CAG GTC CTC AGT CTT TCC TGT TCT GGC ACC TTC ATC AAT	240
ATC ATA GTA ATG TAT TTT GTT GGT GCT CTA TTT GGT GTT TTT CCC ATC	288
TCA GGA ATC CTT TTC TCT TAC TAT AAA ATA GTT TCC ACT ATT CTG AGA	336
ATC CCA TCC TCA GGT GGG AAA TAT AAA GCC TTC TCT ACC TGT GGG TCT	384
CAC CTA TCA GTT GTT TGT TTA TTT TAT GGA ACA GGC TTT GGA GTG TAC	432
CTT GGT TCA GCT GTG TCA CAT TCT TCT AGA AAA TCT GCA GTG GCC TCG	480
GTG ATG TAC ACA GTT GTC ACC CCC ATG CTC AAC CCC TTC ATC TGC AGT	528

(37) INFORMATION POUR LA SEQ ID NO:36 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 504
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 13

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:36 :

GGG GCC ATT CGC TTT CCC CTG CAC AAT ACT ACC ATC ATG AGC CCC AAG	48
CTC GGT CTC TTC CTG GTG GTG CTG TCC TGG GTG CTA ACC ATG TTC CAT	96
GCC ATG CTC CAT ACC CTG CTT ATG GCC AGA TTG TGT TTC TGT GCA GAG	144
AAC ATG ATT CCC CAT TTT TTC TGT GAT ATG TCT GCC CTT CTG AAG CTG	192
TCC TGC TCC AAC ACT CAT GTC AAT GAG TTG GTG ATA TTC ATC ACA GCA	240
GGC CTC ATT CTT CTC ATT CCA TTG GTC CTC ATT CTT CTT TCC TAT GGG	288
CAC ATC GTG TCC TCC ATT CTC AAG GTC CCT TCT GCT CGA GGT ATC CAT	336
AAG ACC TTC TCC ACC TGT GGC TCC CAT TTG TCT GTG GTG TCA CTG TTC	384
TAT GGG ACA GTC ATC GGA CTC TAC TTA TGT CCA TCA GCT AAT AAT TCT	432
ACT GTG AAA GAT ACT GTC ATG GCT CTG ATG TAC ACG GTG GTC ACT CCC	480
ATG CTC AAT CCC TTT ATC TGC AGT	504

(38) INFORMATION POUR LA SEQ ID NO:37 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 528
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 14

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:37 :

CAA GCT TTG GCG TAT GAC AGA TTC CTG GCC ATA TGT CAC CCA CTG CAC	48
TAC ACT GCC ATC ATG AAT CCC AGG CTC TGT GGT TTG CTG GTT CTG GTG	96

TGC TGG ATC CTG AGT GTC CTG CAT GCC TTG TTG CAA AGC TTA ATG GTG	144
TTG CGA CTG TCC TTC TGC AGA GAC ATA GAA ATC CCC CAT TTT TTC TGT	192
GAA CTC AAC CAG GTG GTC CAA CTT GCC TGT TTT GAC AAC CTT CTT AAT	240
GAC ATA GTG ATG AAT TTT GCA CTT GTG CTC TTG GCT ACT TGT CCC CTC	288
GCT GGC ATT CTT TAC TCC TAC TCC AAG ATA GTC TCC TCC ATC CGT GCA	336
ATC TCT TCA GCT CAG GGC AAG TAC AAA GCC TTT TCC ACC TGT GCC TCC	384
CAC CTC TCT GTG GTC TCC TTA TTT TAC TGC ACA GGC CTG GGT GTG TAC	432
CTC AGT TCT GCT GTA TCC CAC AGC TCA CGC TCC AGT GCA ACA GCC TCA	480
GTG ATG TAC ACC GTG GTC ACC CCC ATG CTC AAC CCC TTC ATC TGC AGT	528

(39) INFORMATION POUR LA SEQ ID NO:38 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 357
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 15

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:38 :

CAC CTT TGC AGG TTG CAT CTC ACA GTC CTC AAG CTC GCC TGC TCT GAC	48
ACC CTC ATC AAC AAC ATA GTG GTG TTC TCT ATG ATC ATC GTC CTG GGT	96
GTC TTC CCT CTC AGT GGC ATC CTC TTC TCC TAC TCT CAG ATT TTC TCC	144
TCC ATC CTG AGG ATC TCA TCA GAC AGA GGC AAG TAC AAA GTC TTC TCC	192
ACC TGT GGG TCT CAC CTC CTG GTG GTC TCC TTG TTC TAT GGC AGT AGC	240
CTT GGG GTC TAC CTC AGT TCT GTA GCC ACA CTG TCT TCT AGG ATG ACT	288
CTG ATG GCC TCA GTG ATG TAC ACC ATG GTC ACC CCC ATG CTG AAC CCC	336
ATC ATC TAC ACC CTC CGG AAC	357

(40) INFORMATION POUR LA SEQ ID NO:39 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 477
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 16

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:39 :

CCC ATG TAC TTG TTC CTC GGT AAT TTG TCC TTC CTG GAG ATC CTT TAT	48
ACA TCC ACA GTG GTG CCG AAA ATG CTG GAG GGC TTC CTG CAG GTG GCA	96
GCC ATC TCT GTG ACT GGT TGC TTG ACC CAG TTC TTC ATC TTT GGT TCT	144
CTA GCC ACA GCA GAA TGC TTC CTA CTG GCT GTT ATG GCA TAT GAT CGC	192
TTC TTG GCA ATC TGC TAC CCA CTT CGC TAT CCA CTC CTG ATG GGG CCT	240
AGA TGG TGC ATG GGG CTG GTG GTC ACA GCC TGG CTG TCT GGC TTC ATG	288
GTA GAT GAA TTA GTT GTG GTC CTG ATG GCC CAG CTG AGG TTC TGT GGC	336
TCC AAT CGC ATC GAT CAC TTT TAC TGT CAC TTC ATG CCT TTG GTG GTC	384

CTG GCT TGC TCA GAT CCC CGA GTA GCC CAG GTG ACA ACA TTT GTT CTC	432
TCT GTA GTC CCC CTC ACT GTT CCA TTC GGA CTG ATT CTG ACA TCC	477

(41) INFORMATION POUR LA SEQ ID NO:40 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 339
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 17

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:40 :

GAG GAT CTA TGT GCG AGA CTC AAG CGA TCC AGG TCG GAC ACC ACC ATC	48
AAT GAG GTG GGT ATT CTC ACG GCA GGA TCA GCA GTT GTT ACC CTG CCA	96
TTC ATG TGC ATT CTG GTC TCA TAT GGC CAC ATG GGG GCC ACC ATC CTG	144
AGA AGA CCC TCC CTC AAG GGC ATG TGC AAA GCC TTA TCC ACA TGT GGC	192
TCC CAC CTC TGT GTG GTC TCT GTG TAC TAT GGA GCA GTT ATT GCA CTC	240
TAT ATT GTC CCC TCA TCT AAT AGC ACT AAT GAC AAG GAT ATT GCT GTG	288
TCT GTG TTG TAT ACT CTG GTC ATC CCC ATG CTC AAC CCC TTC ATC TGC	336
AGT	339

(42) INFORMATION POUR LA SEQ ID NO:41 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 528
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 18

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:41 :

CAA GCT TTG GGG TAT GAT AGA TTT GTG GCC ATG TGT CAT CCT GGG CAG	48
TAT CTG GTC ATT ATG AGC CCT CGC CAT GGT GGC TTC CTA ACT TTG GTG	96
TCA TTT TTG CTG AGT CTT TTG GAC TCC CAG CTG CAC AGT TTC ATG ACC	144
TTA AAT ATT ACC AGC TTC AAG GAT GTG GAA ATT TCT AAT TTC TTC TGT	192
GAC CCT TCT CAA CTG CTG AAT CTC TCC TGT TCC AAC ACC TTC TCT GAT	240
AAC ATT GTC AAG TAT TTT CTG GGA GCC TTC TAT GGC CTT TTT CCC ATC	288
TCA GGG ATC CTT TTC TCT TAC TAC AAA ATT ATT TCC TCC ATT CTG AGG	336
ATC CCC TCC TTA GGT GGG AAG TAC AAA GCC TTC TCC ACC TGT GGG TCT	384
CAC CTG GCA GTT GTT TGC TTA TTT TTA GTG ACA GCC TCC ACA GTG TAC	432
CTT GGA TCA GTT GCA TCA CAT TCT CCC AGA AAT GAT GTG GTG GCT TCT	480
CTG ATG TAC ACT GTG GTC ACC CCC ATG CTC AAC CCC TTC ATC TGC AGT	528

(43) INFORMATION POUR LA SEQ ID NO:42 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 528
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 19

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:42 :

CAA GCT TTG GCG TAT GAC AGA TTT CTG GCC ATC TGT CAT CCT CTG CAT	48
TAT CTG GTC ATT ATG AGC CCT CGC CAT TGT GGC TTC TTA ACT TTG GTG	96
TCA TTT TTG CTG AGT CTT TTG GAC TCC CAG CTG CAC AGT TTC ATG ACC	144
TTA AAT ATT ACC AGC TTC AAG GAT GTG GAA ATT TCT AAT TTC TTC TGT	192
GAC CCT TCT CAA CTG CTG AAT CTC TCC TGT TCC AAC ACC TTC TCT GAT	240
AAC ATT GTC AAG TAT TTT CTG GGA GCC TTC TAT GGC CTT TTT CCC ATC	288
TCA GGG ATC CTT TTC TCT TAC TAC AAA ATT ATT TCC TCC ATT CTG AGG	336
ATC CCC TCC TTA GGT GGG AAG TAC AAA GCC TTC TCC ACC TGT GGG TCT	384
CAC CTG GCA GTT GTC TGC TTA TTT TTA GTG ACA GCC TCC ACA GTG TAC	432
CTT GGA TCA GTT GCA TCA CAT TCT CCC AGA AAT GAT GTG GTG GCT TCT	480
CTG ATG TAC ACT GTG GTC ACC CCC ATG CTC AAC CCC TTT ATC TGC AGT	528

(44) INFORMATION POUR LA SEQ ID NO:43 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 528
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN

(ix) CARACTERISTIQUES

- (A) NOM/CLE: AMOR 20

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:43 :

CAA GCT TTG GCG TAT GAC AGG TTC CTG GCC ATC TGT CAT CCT CGG CAT	48
TAT CTG GTC ATT ATG AGC CCT CGC CAT TGT GGC TTC TTA ACT TTG GTG	96
TCA TTT TTG CTG AGT CTT TTG GAC TCC CAG CTG CAC AGT TTC ATG ACC	144
TTA AAT ATT ACC AGC TTC AAG GAT GTG GAA ATT TCT AAT TTC TTC TGT	192
GAC CCT TCT CAA CTG CTG AAT CTC TCC TGT TCC AAC ACC TTC TCT GAT	240
AAC ATT GTC AAG TAT TTT CTG GGA GCC TTC TAT GGC CTT TTT CCC ATC	288
TCA GGG ATC CTT TTC TCT TAC TAC AAA ATT ATT TCC TCC ATT CTG AGG	336
ATC CCC TCC TTA GGT GGG AAG TAC AAA GCC TTC TCC ACC TGT GGG TCT	384
CAC CTG GCA GTT GTT TGC TTA TTT TTA GTG ACA GCC TCC ACA GTG TAC	432
CTT GGA TCA GTT GCA TCA CAT TCT CCC AGA AAT GAT GTG GTG GCT TCT	480
CTG ATG TAC ACT GTG GTC ACC CCC ATG CTC AAT CCC TTC ATC TGC AGT	528

(45) INFORMATION POUR LA SEQ ID NO:44 :

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 336
- (B) TYPE: acide nucléique
- (C) NOMBRE DE BRIN: double
- (D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADN
- (ix) CARACTERISTIQUES
 - (A) NOM/CLE: AMOR 21

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:44 :

CCC ATG TAT TTG TTT CTC GGT AAC CTG TCC TTC ATG GAC ATC TGC TTC	48
ACA ACA GTC GTT GTG CCC AAG ATG CTG GCG AAT TTG CTG TCA GAG ACA	96
AAG GGC ATC TCC TAT GTA GGC TGC CTG GTC CAG ATG TAT TTC TTC ATG	144
GCC TTT GGG AAC ACT GAT AGT TAC CTG CTG GCC TCC ATG GCC ATC GAC	192
CGG CTG GTG GCC ATC TGC AAC CCC TTG CAC TAT GAT GTG GCC ATG CGC	240
CCA CAC CGC TGC CTC ATG CTG CTG GGT TCT TGC ACC ATC TCC CAC	288
CTG CAC GCC CTC TTC CGG GTG CTA CTC ATG TCT CGC CTC TCT TTC TGT	336

(46) INFORMATION POUR LA SEQ ID NO:45 :

- (i) CARACTERISTIQUES DE LA SEQUENCE:
 - (A) LONGUEUR: 357
 - (B) TYPE: acide nucléique
 - (C) NOMBRE DE BRIN: double
 - (D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADN
- (ix) CARACTERISTIQUES
 - (A) NOM/CLE: AMOR 22

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:45 :

CAC CTT TGC AGG TTG CAT CTC ACA GTC CTC AAG CTC GCC TGC TCT GAC	48
ACC CTC ATC AAC AAC ATA GTG GTG TTC TCT ATG ATC ATC GTC CTG GGT	96
GTC TTC CCT CTC AGT GGC ATC CTC TTC TCC TAC TCT CAG ATT TTC TCC	144
TCC ATC CTG AGG ATC TCA TCA GAC AGA GGC AAG TAC AAA GTC TTC TCC	192
ACC TGT GGG TCT CAC CTC CTG GTG GTC TCC TTG TTC TAT GGC AGT AGC	240
CTT GGG GTC TAC CTC AGT TCT GTA GCC ACA CTG TCT TCT AGG ATG ACT	288
CTG ATG GCC TCA GTG ATG TAC ACC ATG GTC ACC CCC ATG CTG AAC CCC	336
ATT ATC TAC ACC CTC CGG AAC	357

(47) INFORMATION POUR LA SEQ ID NO:46 :

- (i) CARACTERISTIQUES DE LA SEQUENCE:
 - (A) LONGUEUR: 423
 - (B) TYPE: acide nucléique
 - (C) NOMBRE DE BRIN: double
 - (D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADN
- (ix) CARACTERISTIQUES
 - (A) NOM/CLE: AMOR 23

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:46 :

TGG AGT CTT TTG GAG TCC CAG CTG CAC AGT TTG AGG ACC TTA AAT ATG	48
ACC AGC TTC AGG GAT GTG GAA AGT TCT AAT TTG TTG TGT GAC CCT TCT	96
CAA ATG CTG AAT CTC TCC TGT TCC AAC ACC TTC TCT GAT AAC ATT GTC	144
AAG TAT TTT CTG GGA GCC TTC TAT GGC CTT TTT CCC ATC TCA GGG ATC	192
CTT TTC TCT TAC TAC AAA ATT ATT TCC TCC ATT CTG AGG ATC CCC TCC	240
TTA GGT GGG AAG TAC AAA GCC TTC TCC ACC TGT GGG TCT CAC CTG GCA	288
GTT GTT TGC TTA TTT TTA GTG ACA GCC TCC ACA GTG TAC CTT GGA TCA	336
GTT GCA TCA CAT TCT CCC AGA AAT GAT GTG GTG GCT TCT CTG ATG TAC	384
ACT GTG GTC ACC CCC ATG CTC AAC CCC TTT ATC TGC AGT	423